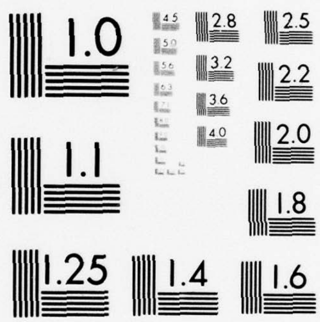
AD-A062 797   MINNESOTA UNIV MINNEAPOLIS DEPT OF PSYCHOLOGY   F/G 5/10
A COMPARISON OF LEVELS AND DIMENSIONS OF PERFORMANCE IN BLACK A--ETC(U)
OCT 78 A T CHURCH, S M PINE, D J WEISS   N00014-76-C-0244
UNCLASSIFIED   RR-78-3   NL

| OF |

AD
A062797

END
DATE
FILMED
3-79
DDC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

LEVEL

(12)

AD A062797

A COMPARISON OF LEVELS AND DIMENSIONS OF PERFORMANCE IN BLACK AND WHITE GROUPS ON TESTS OF VOCABULARY, MATHEMATICS, AND SPATIAL ABILITY

Austin T. Church
Steven M. Pine
and
David J. Weiss

RESEARCH REPORT 78-3
OCTOBER 1978

PSYCHOMETRIC METHODS PROGRAM
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF MINNESOTA
MINNEAPOLIS, MN 55455

Prepared under contract No. N00014-76-C-0244, NR150-383
with the Personnel and Training Research Programs
Psychological Sciences Division
Office of Naval Research

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
9 Research Report 78-3

2. GOVT ACCESSION NO.
14 RR-78-3

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
6 A Comparison of Levels and Dimensions of Performance in Black and White Groups on Tests of Vocabulary, Mathematics, and Spatial Ability.

5. TYPE OF REPORT & PERIOD COVERED
Technical Report

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
10 Austin T./Church, Steven M./Pine[redacted]
David J./Weiss

8. CONTRACT OR GRANT NUMBER(s)
15 N00014-76-C-0244

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Department of Psychology
University of Minnesota
Minneapolis, Minnesota 55455
61153

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
P.E.:6115N PROJ.:RR042-04
T.A.:RR042-04-01
W.U.:NR150-383

11. CONTROLLING OFFICE NAME AND ADDRESS
Personnel and Training Research Programs
Office of Naval Research
Arlington, Virginia 22217

12. REPORT DATE
11 Oct[redacted]978

13. NUMBER OF PAGES
42

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)
12 51p.

15. SECURITY CLASS. (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

16 RR04204 17 RR0420401

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

ability tests	cultural bias	item difficulty
bias	fairness	item discrimination
item bias	race	item characteristic curve theory
test bias	computerized adaptive testing	item by race interactions
racial bias	item calibration	factor composition

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report examines the nature and extent of ability test performance differences between Black and White high school students on vocabulary, mathematics, and spatial ability tests. Mean differences on total test scores were found for all three tests with Whites averaging higher than Blacks. In the vocabulary test, however, this effect could not be interpreted independently of sex and parents' educational level. Parents' educational levels were significantly related to performance on the vocabulary and spatial tests; in the vocabulary test parental education interacted with the

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

406 024

78 12 27 012

race and sex variables. Separate factor analyses were performed for the Black and White groups to determine the number and nature of dimensions underlying performance for each group. While the number of factors needed to account for the common item variance in each test was the same for Blacks and Whites, items defining each factor and the correlations of factors across the three tests indicated that the nature of the factors was different for the two groups. For the vocabulary test, degree of item bias was evaluated in terms of the difference in item difficulties for Blacks and Whites as indexed by the difficulty (b) parameter of Item Characteristic Curve (ICC) theory. Comparison of the ICC item parameters for the Blacks and Whites showed differences in both difficulties and discriminations. By comparing the index of item bias with the vocabulary factor structures in both groups, a "bias" factor defined by "Black-type" words was identified in the White group. Analysis of racial group differences in relationships among subtest scores and factor scores showed that Whites had more common variance among subtests than Blacks, with the largest differences occurring where the vocabulary test was involved. It was concluded that when the factor structures underlying ability tests differ sufficiently for two or more racial groups, the meaning of mean group performance differences becomes less clear. Investigation of the fairness of psychometric tests should include examination of possible bias at both the item and factor level.

CONTENTS

Acknowledgements

The authors wish to thank Harvy Linder for writing several of the items used in this study and Frederic Lord of Educational Testing Service for providing a number of the vocabulary test items.

Technical Editor: Barbara Leslie Camm

A Comparison of Levels and Dimensions of Performance in Black and White Groups on Tests of Vocabulary, Mathematics, and Spatial Ability

An important current concern of users of ability tests in practical applications is the general trend for some minority groups to obtain significantly lower scores on such tests. For example, in a summary of the individual and group tests reported in the literature for urban children in the North and South, Shuey (1958) noted that the average IQ scores for Black and White adults tested during both world wars have consistently differed by 10 to 20 points and that typically, less than 25% of the Black group has scored above the median for Whites.

The presence of Black-White differences on ability tests seems to be well documented (see also Baughman & Dahlstrom, 1968; Jensen, 1973; Kennedy, Van de Riet, & White, 1963; Loehlin, Lindzey, & Spuhler, 1975; Terman & Merrill, 1960). The question which most recent research has tried to answer is, What is the cause or nature of these differences? One possibility which has been considered is that the tests being used for such applications as educational and industrial selection do not measure the same variables for the different groups. Thus, for example, there may be subsets of items or underlying factors of performance which are different for Blacks and Whites. Depending on the nature of such differences, a cultural bias may or may not be reflected in the test or in certain items in the test.

An approach that has been used to investigate this question is the examination of the relative difficulties of items for different groups, under the assumption that even if mean differences in total test performance exist, the relative difficulties of the items should be the same in the different groups. This approach is usually operationalized by using an analysis of variance to look for significant item-by-race interactions where the difficulty of the items is indexed by proportion-correct score (Cardall & Coffman, 1964; Cleary & Hilton, 1968). Because the proportion-correct score can range only from 0 to 1, differences between two groups may become truncated as the average proportion-correct score of the groups approaches 0 or 1. To avoid this problem, item-by-group interaction studies have sometimes converted proportion-correct indices to normal deviates or to the delta index (Angoff & Ford, 1971; Breland, Stocking, Pinchak, & Abrams, 1974; Gitlitz & Kaufman, 1972).

Angoff and Ford (1971) have developed an elliptical graphing technique using the delta transformations for testing the significance of item-by-group interactions which has been applied to data from the National Longitudinal Study of the Class of 1972 by Breland et al. (1974). This study not only identified items which had significant interactions with race but attempted to interpret the nature of possible bias in the items. Breland et al. found, for example, that the most biased items were those reflecting linguistic differences, primarily between Spanish-speaking groups and others. Mathematics items were found to be easier or more difficult for minorities, depending on how much they reflected "street sense," e.g., items dealing with money. Other attempts to

identify and interpret item-by-group effects include comparing the rank orders of item difficulties between groups and examining the frequency with which groups choose specific multiple-choice distractors in order to gain insight into how the items operate differently in the various groups (e.g., Bartlett, Newbrough, & Tulkin, 1972; Jensen, 1974).

All these methods are, in fact, attempts to show that the test as a whole or certain items within it operate differently for different groups (i.e., measure different dimensions of performance). Perhaps the most direct way, then, to determine whether or not this is the case is to factor analyze the item intercorrelation matrix for each group separately and to examine directly the number and nature of dimensions underlying test performance. If different groups require a different number of factors to explain test variance or if the patterns of item loadings suggest that the nature of the underlying factors in the groups is different, these results could be interpreted as evidence that the test emphasizes different dimensions of performance in different groups.

In general, studies which have looked at comparative factor structures in this way (Flaugher & Rick, 1972; Humphreys & Taber, 1973; Johnson, 1969; Lesser, Fifer, & Clark, 1965; Stodolsky & Lesser, 1967) have concluded that most ability tests measure essentially the same underlying dimensions in all groups. However, the patterns of ability levels on these dimensions do appear to vary from group to group. To the extent that a given test emphasizes particular dimensions of performance, these results suggest that some groups can be expected to be at a disadvantage in terms of their test scores.

To investigate a test in a thorough manner to see if the test or certain items within it measure different dimensions of performance for different groups, some method of assessing item-by-group interaction should be combined with an examination of the comparative factor structures of the test. Although a factor analysis of test items seems the most direct way to investigate differential underlying dimensions for different groups, there may be items in the test that function differently for different groups which should be identified and perhaps eliminated from the test but which are not numerous enough to define an important factor. Thus, the factor analytic approach should be supplemented by some method for studying item-by-group interaction.

Analysis of variance methods for identifying item-by-group interactions are direct ways of isolating specific items which operate differently in different groups, but there are problems with these procedures. For example, such an interaction model is not capable of detecting biased items in those cases in which all (or most) of the test items have roughly the same degree of bias. Secondly, the method provides no objective way of determining whether items isolated by such methods operate differently in different groups for the same reason or for reasons specific to each item. Use of an item-by-group interaction method in conjunction with a factor analysis allows examination of the loadings of the items identified to determine whether or not they define a consistent spurious dimension. Combined with an examination of item content, these analyses can lend extra confidence to interpretations of why the items behave differently for the different groups.

This report is concerned with (1) the nature and extent of test performance differences between Blacks and Whites on three kinds of ability tests and

variables related to those differences; (2) the use of factor analysis to compare the structures of tests for Blacks and Whites; (3) the use of an index of item-by-group interaction derived from the parameters of Item Characteristic Curve (ICC) theory (Lord & Novick, 1968) to detect possibly biased items. These studies are preliminary to later studies evaluating the use of computerized adaptive testing techniques and indices of item bias to reduce differences in levels and dimensions of ability test performance for different groups.

METHOD

Tests

Three tests were constructed to measure vocabulary, mathematics, and spatial abilities. All three tests consisted of a number of five-alternative multiple-choice items printed in a conventional test booklet format.

Vocabulary Test

The vocabulary test consisted of 127 items drawn from several sources. Sixty-five items were drawn from the Black literature sources listed in Appendix Table A-1; eight additional items were developed specifically for this test using "Black-type" words provided by a Black clinical psychologist. Although words were drawn from all sources listed in Appendix Table A-1, only those words drawn from the Dictionary of Afro-American Slang and those provided by the Black clinical psychologist were clearly distinguishable as "Black-type" words. Consequently, where performance comparisons were made for racial groups on "Black-type" words versus "non-Black-type" words, only the 32 words from these two sources were used. Stems for 17 items were drawn from Webster's Seventh New Collegiate Dictionary and represented words which were assumed to be more familiar to Whites. Thus, there were approximately twice as many words in the test written specifically for Blacks as words written specifically for Whites. The remaining items came from already existing items in the University of Minnesota Psychometric Methods adaptive testing item pool (McBride & Weiss, 1974) and from items provided by Educational Testing Service. These items are typical of vocabulary items administered in standardized tests, such as the Scholastic Aptitude Test. Examples of items from all these sources are shown in Appendix Table A-2.

Mathematics Test

The mathematics test consisted of 25 multiple-choice items involving application of basic arithmetic operations (singly or in combination), simple unit conversions, and "story" problems involving quantitative reasoning. All items were selected from a locally developed experimental item pool and were similar to items in the mathematics subtest of the General Aptitude Test Battery (GATB). Sample items are contained in Appendix Table A-3.

Spatial Test

The spatial test consisted of 30 items which required the subject to visualize which of 4 three-dimensional solids would result from "folding" the one-dimensional pattern in the stem. Similar to the mathematics items, items

were selected from a locally developed pool and were similar to items in the spatial subtest of the GATB. Appendix Figure B-1 contains some sample spatial items.

Procedure

The vocabulary, mathematics, and spatial tests were given in two administrations to students at two racially mixed high schools in Minneapolis. The vocabulary test was given during the first testing, and the mathematics and spatial tests were combined into one booklet for the second administration. There was one day between administration of the vocabulary test and the mathematics and spatial tests in both schools. The testees were 10th through 12th grade students at one school and 9th grade students at the other. The vocabulary test was given to a total of 285 students, including 92 Blacks, 173 Whites, 4 Latin Americans, 4 Native Americans, and 12 for whom no ethnic information was reported. Half the vocabulary test booklets contained the pages in reverse order so that data would be obtained on all items in case a large number of students were not able to complete the test. It was found that all students were able to complete the tests in the time allotted.

The mathematics and spatial tests were completed by 300 students, including 86 Blacks, 183 Whites, 2 Latin Americans, 7 Spanish Americans, and 22 who did not provide information on ethnic background. A total of 161 students were present at both test administrations at their respective schools so that scores for all three tests were available for this subset of the total group. In addition to the tests, all students completed a biographical questionnaire at each testing session providing information on their sex, race, and the highest educational level obtained by their mothers and fathers. A copy of this biographical form appears in Appendix Figure B-2.

Total testing time allowed for the 127-item vocabulary test was 60 minutes, and 60 minutes were also allotted for the combined mathematics and spatial tests of 25 and 30 items, respectively. At one school both test administrations were proctored by six teachers under the supervision of the chief counselor and a mathematics teacher. At the other school the chief counselor, two teachers, and two psychology research assistants served as proctors. An identical set of test instructions was used and read at all test administrations.

Data Analysis

Mean Differences

The effects of race, parents' educational level, and sex on test scores were analyzed separately for each of the three tests using a three-way analysis of variance. Because cell frequencies in the three-way crossed classification were neither exactly equal nor proportional, analysis of variance computational procedures which account for nonorthogonal main effects and nonorthogonality between main and interaction effects were used. Computations were based on the "classic experimental" approach described by Nie, Hull, Jenkins, Steinbrenner, and Bent (1975, pp. 405-408).

An initial descriptive index of the educational level for each parent was obtained by assigning a number from 1 to 6 to each of the six educational

response categories in the biographical questionnaire as follows: 1 = not applicable, 2 = did not complete high school, 3 = finished high school, 4 = some college or other post-high school, 5 = finished college, 6 = attended graduate or professional school. Using these numerical values, mean educational level for fathers and mothers of both racial groups was then computed for those students who reported Categories 2 through 6. The frequency of responses to each educational category for both the vocabulary and mathematics/spatial administrations is shown in Appendix Table A-4. Responses for students who scored below chance levels (see below) were not included in these frequencies.

In order to obtain a single index of parental education for each student, the higher of the two reported educational levels for father and mother was used. Since the sample sizes at some educational levels were rather small for analysis of variance purposes, this single index of parental education was then collapsed into three groups, depending on whether the higher educated parent (1) did not complete high school, (2) did complete high school or the equivalent, or (3) had at least some college or additional schooling after high school. This latter group also included those families in which one or more parents had actually completed college or had attended graduate or professional school. In assigning students to educational levels, it was assumed that those who did not respond to the question about father's or mother's education or answered "not applicable" did not have either a father or a mother, or both, living in the home. Twenty-four students did not provide parents' educational level for the vocabulary test analysis, and 48 students did not provide this information for the mathematics and spatial tests. Assuming that parental environment affects ability test scores, it is reasonable to include these students in the lowest educational level for the analysis. To the extent that test scores of these students averaged lower than the other students included in the lowest educational level (those reporting some high school but no diploma), the effects of parents' educational level on scores will be somewhat overestimated if those not providing parents' educational information did in fact have parents with a high school diploma or some post-high school education or training.

For each test separately, scores below the chance level (defined as the number of items in the test divided by the number of response alternatives for each item) were not included in the calculations. The original 265 vocabulary scores included 6 such chance scores for Whites and 2 for Blacks which could be obtained on the basis of guessing alone; 5 Blacks and 3 Whites were eliminated from the mathematics and spatial test analyses for this reason. Other students eliminated from the analyses were those who did not provide racial information or were either Native American or Latin American.

Factor Composition

For each of the tests, an item intercorrelation matrix of phi coefficients was obtained for both of the racial groups. All items were included in the factor analyses for the mathematics and spatial subtests; due to computer limitations, however, 75 randomly selected items were used in the factor analysis of the vocabulary items. Although there have been arguments in favor of using tetrachoric correlations in such factor analyses because phi coefficients are more likely to result in "difficulty" factors (Lord & Novick, 1968, p. 382), Lumsden (1976) concluded that normally it should not be necessary to use tetrachorics. He cited a study by McDonald and Ahlawat (1974) which

demonstrated that difficulty factors will not arise with phi coefficients if the regressions of the items on the latent trait are linear and that difficulty factors will be negligible unless the item difficulties are widely different or the item discriminations are high. McDonald and Ahlawat concluded that the widely held belief that spurious factors occur when phi coefficients are "attenuated" due to unequal difficulties is false.

For each test within both the Black and White groups, a principal axis solution was obtained, and the important factors were rotated by the varimax procedure. The number of factors extracted and rotated in each case, and the final number of rotated factors scored for each test, were determined through the following process:

1. For each principal axis solution, a parallel analysis (Humphreys & Ilgen, 1969) of random correlation matrices with the same number of subjects and items was conducted, and the eigenvalues of the real data solution were compared to those of the random factor solution. Real data factors with eigenvalues greater than the corresponding random factors were tentatively considered to be meaningful.

2. When the parallel analysis procedure did not give a clear indication of the number of factors to retain, the rotated eigenvalues and the loadings of items on each factor were examined to determine which number of factors better explained the data.

3. In all cases, rotated factor solutions were obtained for one or two factors above and below the final factor solution to determine whether the rotated eigenvalues and item factor loadings were relatively stable and/or whether more or fewer factors would provide a better final solution.

Item Bias Analysis

One purpose of this research was to develop a calibrated pool of vocabulary test items to be used in later research on the reduction of test bias by computerized adaptive testing methods. Because vocabulary tests have generally been shown to be relatively unidimensional, this analysis was confined to the vocabulary items; the assumption of unidimensionality was, of course, tested by the factor analysis of these items. Estimates of the discrimination (a), difficulty (b), and guessing (c) parameters of Item Characteristic Curve (ICC) theory (Lord & Novick, 1968) were obtained using a computer program which implements the shortcut estimation procedure described by Jensema (1976). This procedure involves fixing a value of the guessing parameter (c), which is used along with the dichotomous item responses to obtain estimates of the discrimination and difficulty parameters. The method has been shown to be comparable to iterative methods such as Urry's ESTEM when the sample sizes are not large (Jensema, 1976), as was the case in this study.

Item bias index. ICC parameters were obtained for the vocabulary items using the total combined group and the Black and White groups separately. An index of item bias for the vocabulary items was obtained by computing the difference between the difficulty parameters for the Black and White groups. This index is similar to comparing the proportion correct for an item for two

groups, except that the proportion-correct index of difficulty used in classical test theory is confounded by level of discrimination and guessing effects (Lord & Novick, 1968; Urry, 1974).

Variance Overlap Analysis

A relevant question with regard to the study of race differences in ability tests is whether different kinds of abilities are related to each other in the same ways within different racial groups. Thus, scores on the three subtests were intercorrelated for the two racial groups.

Multivariate redundancy analysis was used to further investigate this question. For each test within each racial group, factor scores were obtained for the rotated factors resulting from the factor analyses within the groups for each test. The redundancies of one set of test factors on another were estimated using multiple regression of the factor scores as described by Stewart and Love (1968) and Weiss (1972). These redundancies provide summary information on the variance overlap between factors in pairs of tests.

RESULTS

Mean Differences

Parents' educational level. Table 1 shows the mean and standard deviation of the reported highest educational level achieved by each student's mother and father for those students providing this information. The data are shown separately for the group completing the vocabulary test and the group completing the mathematics and spatial tests. Table 1 also shows the mean and standard deviation of the composite parental education index used in the analysis of variance.

Table 1
Parents' Highest Educational Level by Racial Group

	Blacks			Whites		
Test and Parent	*N*	Mean	S.D.	*N*	Mean	S.D.
Vocabulary Test						
Father	63	3.39	1.17	150	3.16	1.10
Mother	73	3.60	1.18	160	3.17	.93
Composite Index	90	2.17	.86	167	2.20	.69
Mathematics and Spatial Tests						
Father	50	3.26	1.03	152	3.11	1.08
Mother	52	3.46	1.11	161	3.06	.78
Composite Index	81	1.94	.87	180	2.05	.68

For both the group completing the vocabulary test and the group completing the mathematics and spatial tests, the mean difference between mothers' educational level for Blacks and Whites was statistically significant ($p < .005$); mothers of Black students were reported to have higher educational levels than mothers of White students. Although Blacks reported somewhat higher fathers' educational levels than Whites, the mean difference was not statistically

significant. When the highest level of education attained by either parent was used as a composite parental education index, there was no significant difference between Blacks and Whites. The composite index was slightly lower for Blacks than for Whites in Table 1 because there were more Black students who did not report parents' educational levels. As described earlier, these students were assumed not to have the parent in question living in the home and were included in the lowest category for the composite educational index and analysis of variance. Thus, while Blacks reporting parental educational levels reported slightly higher levels, the larger number of Blacks not reporting this information resulted in the mean "parental contribution to the educational environment of the home" being about the same for the two groups.

Table 2

Means and Standard Deviations of Vocabulary Test Scores for all Combinations of Race, Parents' Educational Level, and Sex, and Results of the Three-Way ANOVA

	Racial Group						Combined		
	Blacks			Whites			Racial Groups		
Group	*N*	Mean	S.D.	*N*	Mean	S.D.	*N*	Mean	S.D.
Some High School									
Male	10	52.40	15.31	15	57.07	17.44	25	55.20	16.45
Female	17	54.35	15.52	11	53.63	10.00	28	54.07	13.41
High School Diploma									
Male	14	54.29	13.37	44	55.16	16.24	58	54.95	15.49
Female	7	55.00	13.00	37	66.40	14.49	44	64.59	14.74
Some Post-High School									
Male	20	51.05	14.68	27	69.74	17.20	47	61.79	18.53
Female	22	64.41	11.01	33	62.67	19.01	55	63.36	16.19
Combined Groups									
Male	44	52.39	14.15	86	60.07	17.83	130	57.47	17.01
Female	46	59.26	13.77	81	63.15	16.40	127	61.74	15.56
Some High School	27	53.63	15.18	26	55.62	14.60	53	54.60	14.78
High School Diploma	21	54.52	12.92	81	60.30	16.37	102	59.10	15.84
Post-High School	42	58.05	14.41	60	65.85	18.41	102	62.64	17.24
Total Group	90	55.90	14.30	167	61.56	17.17	257	59.58	16.42

Three-Way Analysis of Variance

Source of Variation	Sum of Squares	*DF*	Mean Square	*F*	p^a
Main Effects	5225.75	4	1306.44	5.39	.001
Race	1821.66	1	1821.66	7.52	.007
Education	2102.80	2	1051.40	4.34	.014
Sex	1117.00	1	1117.00	4.61	.031
Two-Way Interactions	2159.08	5	431.82	1.78	.116
Race × Education	370.60	2	185.28	.76	.999
Race × Sex	804.79	1	804.79	3.32	.066
Education × Sex	1616.92	2	808.46	3.34	.036
Three-Way Interaction					
Race × Education × Sex	2271.70	2	1135.78	4.69	.010
Residual	59358.21	245	242.28		
Total	69014.62	256	269.59		

[a] Estimated probability of error in rejecting null hypothesis.

The data suggest that the sample was drawn from a population which contained parents of all educational levels, ranging from having had some high school to having attended graduate or professional school. The average parent of both races had completed a high school diploma and in some cases (especially mothers of Black students) was likely to have had some schooling beyond high school, e.g., some college or trade school.

Vocabulary test scores. Table 2 shows the means and standard deviations of vocabulary number-correct scores for all combinations of race, the three parental educational levels, and sex, and for all subgroups combined, as well as the results of the three-way analysis of variance. The main effects for race ($p<.01$), parents' educational level ($p<.05$), and sex ($p<.05$) were all statistically significant. Mean scores of Whites (mean=61.56) were significantly higher as a group on the vocabulary test than those of Blacks (mean=55.90). Parents' educational level was related to mean vocabulary scores, with students having more highly educated parents scoring higher on the average. Females (mean=61.73) had average scores which were higher than those of males (mean=57.47) across the total sample. These significant main effects must be qualified, however, since there was a significant ($p<.05$) two-way interaction between parents' educational level and sex and a significant three-way interaction between race, parents' educational level, and sex ($p<.01$).

Figure 1
Three-Way Interaction Among Race, Parents' Educational Level, and Sex for Vocabulary Test Scores

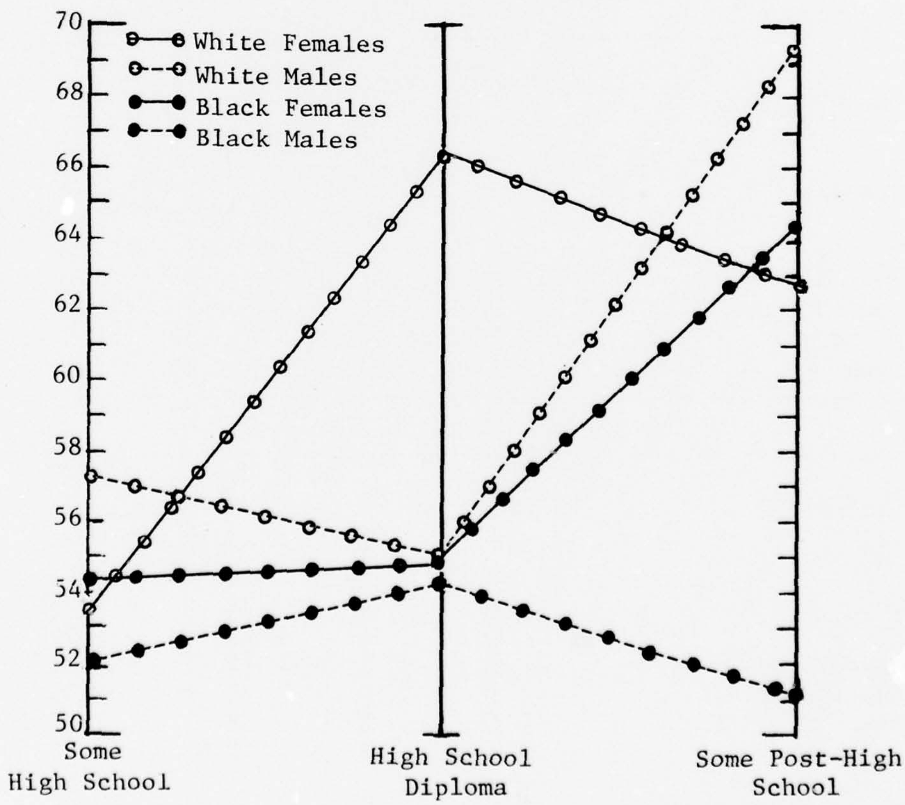


Figure 1 permits examination of the significant interactions. The departure from a consistent trend across educational levels was determined for the most part by the Black males, whose mean performance was lowest at the highest parental educational level, and the White females, whose mean performance was highest at the second educational level. Thus, the distinction between "some high school" and "high school diploma" seemed important only for the White females in this sample, while the distinction between "high school diploma" and "some post-high school" seemed important for all groups except the Black males. That no parental measure of general or verbal intelligence was obtained in this study, and the rather broad nature of the highest educational category used, makes interpretation of the interactions found difficult and speculative. Most important is the demonstration that the main effect for race found in this test can not be interpreted independently of the sex of the students and their parents' educational levels.

Table 3

Means and Standard Deviations of Mathematics Test Scores for all Combinations of Race, Parents' Educational Level, and Sex, and Results of the Three-Way ANOVA

	Racial Group						Combined Racial Groups		
	Blacks			Whites					
Group	*N*	Mean	S.D.	*N*	Mean	S.D.	*N*	Mean	S.D.
Some High School									
Male	16	10.81	2.81	15	13.27	3.80	31	12.00	3.50
Female	16	10.69	3.87	22	12.04	2.98	38	11.47	3.41
High School Diploma									
Male	13	10.46	3.59	47	13.59	4.56	60	12.92	4.53
Female	7	11.86	2.48	49	14.28	3.92	56	13.98	3.84
Some Post-High School									
Male	11	9.27	2.41	18	14.28	4.61	29	12.38	4.59
Female	15	11.60	4.27	29	14.93	4.22	44	13.80	4.48
Combined Groups									
Male	40	10.28	2.99	80	13.69	4.40	120	12.55	4.29
Female	38	11.26	3.76	100	13.98	3.94	138	13.23	4.07
Some High School	32	10.61	3.38	37	12.54	3.35	69	11.63	3.48
High School Diploma	20	10.95	3.25	96	13.95	4.24	116	13.43	4.23
Post-High School	26	10.75	3.63	47	14.68	4.33	73	13.21	4.49
Total Group	81	10.74	3.39	180	13.85	4.14	261	12.88	4.18

Three-Way Analysis of Variance

Source of Variation	Sum of Squares	*DF*	Mean Square	*F*	p^a
Main Effects	591.26	4	147.82	9.58	.001
Race	418.60	1	418.60	27.14	.001
Education	54.19	2	27.09	1.76	.173
Sex	16.32	1	16.32	1.06	.305
Two-Way Interactions	98.56	5	19.71	1.28	.273
Race × Education	41.96	2	20.98	1.36	.257
Race × Sex	16.42	1	16.42	1.06	.304
Education × Sex	45.73	2	22.87	1.48	.227
Three-Way Interaction					
Race × Education × Sex	1.87	2	.94	.06	.999
Residual	3794.42	246	15.42		
Total	4486.12	257	17.46		

[a]Estimated probability of error in rejecting null hypothesis.

Mathematics test scores. Table 3 shows the means and standard deviations of number-correct scores on the mathematics test for all combinations of race, parental education, and sex, and the results of the three-way analysis of variance. Since sex was not reported or determined for three Black students, the cell means and analysis of variance were based on 78 of the original 81 Blacks with non-chance mathematics scores. For the mathematics test, only the main effect for race ($p<.001$) was statistically significant. White students obtained significantly higher mean scores on this test (mean=13.85) than did Black students (mean=10.74). There were no significant two-way interactions, nor was the three-way interaction statistically significant for this test.

It is of interest that the educational index of the parents was not related to scores on the mathematics test and that there was no sex difference in mathematics scores, as is often found in tests of quantitative skills. One possible explanation for this result may be found in the low mean scores on the mathematics test across all groups, which might indicate that the mathematics test was uniformly too difficult for the students tested.

Spatial test scores. Means and standard deviations of spatial test number-correct scores for all groups are shown in Table 4, along with the results of the three-way ANOVA. Since sex was not reported or determined for three Black students, the cell means and analysis of variance were based on 78 of the original 81 Blacks with non-chance spatial scores. For this test, the main effects for both race ($p<.001$) and parents' educational level ($p<.05$) were statistically significant. There were no significant two-way or three-way interactions.

The race effect indicated higher mean scores for the Whites (21.45) than for the Blacks (18.83). Examination of the within-cell means in Table 4 indicates that a parent having completed high school was associated with higher spatial scores for both groups, but that the post-high school distinction was not associated with practically significant differences in mean scores. In fact, lower mean scores were shown for Black females whose parents had some education beyond high school.

Summary. For all three ability tests administered, the White group scored significantly higher than the Black group, although this effect could not be interpreted independently of the sex and parents' educational level of the students on the vocabulary test. This effect cannot be attributed to social class differences, at least as indexed by parental education, since Blacks and Whites were comparable on this variable.

The results for the parental education variable were not as straightforward. In the spatial test, the main effect for parental education was statistically significant, but the effect was not large and was determined primarily by the distinction between the lowest and middle educational levels. In the vocabulary test, scores generally increased with increasing parental education except for the Black males. The trend of the increases over educational levels varied depending on the sex and race of the student, however. No educational effect was found in the mathematics test.

The sex variable was only important for the vocabulary test. Black females scored higher than Black males at all parental educational levels, but the sex variable interacted with educational level for the White students.

Table 4
Means and Standard Deviations of Spatial Test Scores for all Combinations of Race, Parents' Educational Level, and Sex, and Results of the Three-Way ANOVA

	Racial Group						Combined Racial Groups		
	Blacks			Whites					
Group	*N*	Mean	S.D.	*N*	Mean	S.D.	*N*	Mean	S.D.
Some High School									
Male	16	18.56	4.35	15	20.60	5.02	31	19.55	4.72
Female	16	17.62	3.28	22	19.50	4.00	38	18.71	3.80
High School Diploma									
Male	13	19.62	4.89	47	21.68	4.27	60	21.23	4.45
Female	7	20.28	3.35	49	21.82	3.28	56	21.63	3.30
Some Post-High School									
Male	11	19.73	2.87	18	22.00	3.05	29	21.14	3.14
Female	15	18.33	3.13	29	22.03	3.23	44	20.77	3.63
Combined Groups									
Male	40	19.23	4.13	80	21.55	4.16	120	20.78	4.28
Female	38	18.40	3.29	100	21.37	3.55	138	20.55	3.72
Some High School	32	18.27	3.90	37	19.95	4.42	69	19.16	4.23
High School Diploma	20	19.85	4.33	96	21.75	3.78	116	21.42	3.92
Post-High School	26	18.75	3.08	47	22.02	3.13	73	20.80	3.47
Total Group	81	18.83	3.76	180	21.45	3.82	261	20.64	3.99

Three-Way Analysis of Variance

Source of Variation	Sum of Squares	*DF*	Mean Square	*F*	p^a
Main Effects	512.33	4	128.08	8.96	.001
Race	267.94	1	267.94	18.74	.001
Education	126.92	2	63.46	4.44	.013
Sex	6.76	1	6.76	.47	.999
Two-Way Interactions	33.95	5	6.79	.48	.999
Race × Education	15.20	2	7.60	.53	.999
Race × Sex	.76	1	.76	.05	.999
Education × Sex	13.75	2	6.87	.48	.999
Three-Way Interaction					
Race × Education × Sex	8.68	2	4.34	.30	.999
Residual	3517.33	246	14.30		
Total	4072.30	257	15.85		

[a] Estimated probability of error in rejecting null hypothesis.

Factor Composition

Vocabulary test. Table 5 shows the eigenvalues of the unrotated common factor matrices of random and real data for the Black and White groups. A comparison of the unrotated eigenvalues for the first five factors in the Black group with the eigenvalues obtained from the random data suggested that three or four factors represented more than "chance" variation in the data. On the basis of an examination of the rotated factor contributions for the four-factor solution (not shown here) and the item factor loadings on each factor, it was concluded that the three-factor solution adequately accounted for the

non-random common variance in the items. For the White group, the comparison of random versus real eigenvalues, the rotated factor contributions, and the item factor loadings suggested that three factors adequately accounted for the meaningful common variance in the vocabulary test.

Table 5
Eigenvalues of the Unrotated Random and Real Vocabulary Item Intercorrelation Matrices for Blacks and Whites, and Factor Contributions for the First Three Factors of the Rotated Factor Matrices

Group and Factor	Unrotated Solution: Random Data	Unrotated Solution: Real Data	Unrotated Solution: Percent of Common Variance	Unrotated Solution: Percent of Total Variance	Rotated Solution: Factor Contribution	Rotated Solution: Percent of Common Variance	Rotated Solution: Percent of Total Variance
Blacks							
1	2.42	7.63	46.7	10.1	5.69	46.9	7.6
2	2.21	2.44	14.8	3.2	4.00	32.9	5.3
3	2.16	2.23	13.5	2.9	2.45	20.2	3.3
4	2.06	2.08	12.7	2.8	----	----	---
5	2.03	2.01	12.2	2.7	----	----	---
Whites							
1	1.69	11.36	69.0	15.1	6.25	41.9	8.3
2	1.57	1.96	12.0	2.6	5.54	37.2	7.4
3	1.48	1.60	9.8	2.1	3.12	20.9	4.2
4	1.39	1.35	8.3	1.8	----	----	---

Thus, a three-factor solution accounted for the common item variance in both the Black and White groups. The factor contributions and percentages of common and total variance accounted for by each rotated factor in the final three-factor solution are shown in Table 5. Factor loadings for the vocabulary test items are shown in Table 6. Although the number of factors was the same in both racial groups, examination of the factor loadings on similarly ordered factors suggested that the factor composition of the vocabulary test for the two racial groups was not highly similar. In addition, the rotated factor contributions of the three factors were uniformly higher for the White group than for the Black group. This implies more common variance among the items for Whites than Blacks or, conversely, more error or specific variance in the scores of Blacks than Whites.

Table 7 shows the intercorrelations of factor loadings for the three factors of the Black group and those of the White group. Table 7 shows that there was moderate similarity (r=.58) between the second largest White vocabulary factor and the largest Black vocabulary factor. While some of the other correlations were statistically significant, none of the other Black versus White factor pairs seemed to be determined or defined by the same vocabulary items. Thus, the question can be raised as to what the first vocabulary factor in the White group represented, since it did not appear to have a corresponding factor in the Black group.

Table 6
Rotated Factor Loadings of Vocabulary Test Items for Black and White Groups

Item	Blacks Factor 1	Blacks Factor 2	Blacks Factor 3	Blacks h^2	Whites Factor 1	Whites Factor 2	Whites Factor 3	Whites h^2
2	.21	.31	.05	.14	.26	.20	.18	.14
3	.09	.27	.16	.11	.12	.08	.24	.08
5	.32	.05	.01	.10	.18	.17	.04	.06
7	-.03	.29	.60	.45	-.08	-.02	.32	.11
8	.24	.51	-.08	.33	.48	.23	.09	.30
9	.11	.28	-.36	.22	.36	.21	.23	.23
12	.51	.34	-.08	.39	.27	.46	.37	.42
14	.24	.11	-.26	.14	.31	.35	.11	.23
15	-.00	.04	-.14	.02	.02	.06	-.09	.01
16	.05	.49	.13	.26	.42	-.01	.28	.26
17	.25	.10	.19	.11	.39	.39	.04	.31
18	.17	.31	-.07	.13	.19	.19	.33	.19
19	.10	.14	.24	.09	.07	.12	.53	.30
20	.48	.02	-.10	.24	.32	.41	.08	.28
24	.08	.24	-.10	.07	.03	.14	-.18	.05
26	.26	.27	-.12	.16	.22	.11	.36	.10
28	.16	.33	.12	.15	.48	.40	.10	.40
29	.17	.44	.08	.23	.38	.15	.34	.28
31	-.07	.11	-.09	.03	-.05	-.20	-.45	.24
32	.38	.19	.12	.20	.22	.25	.03	.12
33	.25	.14	.27	.15	-.04	.06	.26	.08
34	-.02	.06	-.25	.06	.16	.09	-.02	.03
36	.32	.14	-.03	.12	.12	.39	.17	.19
37	.32	.08	-.13	.13	.37	.24	.16	.22
38	-.09	-.05	.25	.08	.39	.11	.03	.16
39	-.07	.24	.15	.09	.13	.25	.37	.22
41	.36	.02	-.03	.13	.49	.02	.08	.25
44	.14	.12	-.08	.04	.52	.06	.08	.28
45	.42	-.07	-.07	.19	.45	.35	.08	.34
46	-.04	.23	.00	.06	-.08	.13	-.04	.03
48	-.00	.06	-.15	.03	.14	-.28	.25	.16
49	-.11	.30	.15	.12	-.01	.29	.22	.13
50	-.01	.07	.29	.09	.13	.13	.26	.10
52	.18	.01	-.09	.04	.43	.26	.27	.32
53	.30	.20	-.04	.13	.13	.18	.20	.09
59	.24	.07	-.02	.06	-.18	.02	.07	.04
62	-.01	-.01	-.06	.00	.38	.08	.07	.15
65	-.02	.14	.03	.02	.15	.08	-.05	.03
66	.04	.40	.00	.16	.29	.40	.13	.26
67	.01	.02	-.29	.08	.30	-.03	.36	.22
69	.01	.41	.15	.19	.34	.24	.25	.23
71	.09	-.11	-.27	.09	.39	.02	.10	.17
79	.09	.27	.28	.16	.12	-.02	-.02	.01
80	.07	.15	-.16	.05	.39	.14	-.01	.17
81	.22	.11	.20	.10	.13	.46	.26	.30
82	.27	.35	-.00	.19	.13	.31	.10	.12
85	.38	.12	.33	.26	.22	.53	.10	.34
87	.29	.31	-.05	.18	.11	-.04	.08	.02

(continued on next page)

Table 6 (continued)
Rotated Factor Loadings of Vocabulary Test Items for Black and White Groups

	Blacks				Whites			
Item	Factor 1	Factor 2	Factor 3	h^2	Factor 1	Factor 2	Factor 3	h^2
88	.37	.10	.25	.21	.23	.34	.09	.18
90	.29	.13	.18	.13	.41	.26	.19	.27
91	.33	.46	-.19	.35	.27	.31	.27	.25
92	.17	-.03	.00	.03	.60	.22	.16	.44
94	.61	.24	-.07	.43	.12	.43	.15	.23
96	.39	.32	-.01	.25	.21	.48	-.08	.28
98	.23	.24	.34	.23	.25	.39	.05	.21
99	.30	.50	.01	.34	.13	.48	.14	.27
100	-.03	.22	.19	.08	.34	.23	-.11	.18
101	.40	.13	.09	.19	.46	.20	-.01	.25
102	.13	.21	.17	.09	.03	.07	.20	.05
103	.02	.40	-.12	.17	.42	.22	.28	.30
105	.20	.56	-.02	.35	.05	.45	.36	.33
106	.32	-.09	.00	.11	.41	.18	.18	.23
107	.44	-.09	.03	.20	.57	.12	.19	.38
110	.36	-.01	.06	.14	-.01	.17	.17	.06
111	.53	.01	.16	.31	.03	.64	.07	.41
113	.43	-.08	.18	.22	-.01	.32	.06	.11
114	.49	-.06	.41	.41	.30	.35	.06	.21
115	.06	-.21	.15	.07	-.07	-.09	.05	.01
116	.43	.19	.21	.27	.34	.26	.15	.21
117	.25	.01	.25	.13	.26	.26	.03	.14
118	.17	.13	-.08	.05	.34	.16	.33	.25
122	.39	.12	.28	.25	.27	.31	.02	.17
124	.51	.17	.08	.29	.16	.51	.15	.31
125	.46	.21	.20	.30	.17	.29	.07	.12
126	.32	.10	-.10	.12	.39	.14	.19	.21
Factor Contribution	5.69	4.00	2.45	12.23	6.25	5.54	3.12	14.85

Mathematics test. The factor structures of the mathematics test appeared to be more similar for the Black and White groups than were the factor structures of the vocabulary tests. A comparison of random and real eigenvalues

Table 7
Correlations of Rotated Factor Loadings for Vocabulary Test Items

Factor in	Factor in White Group		
Black Group	1	2	3
1	.16	.58***	.02
2	-.04	.20*	.27**
3	-.23*	.16	.05

*$p<.05$; **$p<.01$; ***$p<.001$.

Table 8
Eigenvalues of the Random and Real Mathematics Item Intercorrelation Matrices for Blacks and Whites, and Factor Contributions for the First Three Factors of the Rotated Factor Matrices

	Unrotated Solution				Rotated Solution		
			Percent of		Factor	Percent of	
Group and Factor	Random Data	Real Data	Common Variance	Total Variance	Contri-bution	Common Variance	Total Variance
Blacks							
1	1.36	2.09	42.1	8.4	2.05	53.9	8.2
2	1.30	1.84	37.6	7.4	1.75	46.0	7.0
3	1.06	.96	19.6	3.8	----	----	---
Whites							
1	.90	2.98	65.2	12.0	2.34	60.9	9.4
2	.64	.86	19.3	3.6	1.50	39.1	6.0
3	.72	.71	15.4	2.8	----	----	---

Table 9
Rotated Factor Loadings of the Mathematics Test Items for Black and White Groups

	Blacks			Whites		
Item	Factor 1	Factor 2	h^2	Factor 1	Factor 2	h^2
1	.135	-.627	.411	.029	.097	.010
2	.180	-.473	.256	.186	-.044	.037
3	.007	.078	.006	.006	-.027	.001
4	.237	-.024	.057	.345	.129	.136
5	.198	-.010	.039	.166	-.112	.040
6	.145	-.394	.176	.141	-.120	.034
7	.488	.005	.238	.265	.049	.073
8	.377	-.275	.218	.222	.044	.051
9	.177	-.130	.048	.423	.215	.225
10	.406	-.117	.179	.672	-.099	.461
11	.435	.494	.433	.468	.257	.285
12	.397	-.124	.173	.280	.236	.134
13	.340	.296	.203	.398	.209	.202
14	.249	.186	.097	.447	.219	.248
15	.496	.087	.254	.185	.286	.116
16	-.179	.236	.088	.119	.372	.153
17	.281	.023	.079	.348	.129	.138
18	.441	.037	.196	.373	.146	.160
19	.242	.037	.060	.410	.249	.230
20	.239	.327	.164	.310	.242	.155
21	.076	.353	.130	.313	.231	.151
22	-.230	.324	.158	.084	.314	.106
23	.102	-.007	.010	.105	.453	.216
24	.232	-.096	.063	.098	.643	.423
25	.047	.284	.083	-.185	.138	.053
Factor Contri-bution	2.05	1.75	3.819	2.34	1.50	3.838

(Table 8) and item factor loadings (Table 9) suggested that there were two factors in the mathematics test for both groups. The two factors had similar rotated eigenvalues for both groups, suggesting that the rotated factors were of equal strength in both groups. A further examination of the loadings of items on the first factors for the Black and White groups in Table 9 suggests that these factors were defined by the same items in similar strengths. This was further supported by the correlation (r=.56) between item factor loadings on the first rotated mathematics factor for the two groups, shown in Table 10. The correlation across items between the second factors in both groups (r=.39), as well as an inspection of the loadings in Table 9, suggests a modest similarity for the second factors in the two groups.

Table 10
Correlations of Rotated Factor
Loadings for Mathematics Test Items

Factor in Black Group	Factor in White Group 1	Factor in White Group 2
1	.56**	-.16
2	.14	.39*

*p<.05; **p<.01

Spatial test. A comparison of the factor structures for the Black and White groups on the spatial test gave results somewhat different from either of the other two tests. In the White group, the parallel analysis of random and real factors suggested that four factors represented more than chance variation in the items. An examination of the rotated factor contributions and item factor loadings under both the three- and four-factor solutions indicated that three factors adequately accounted for the common variance in the items. In the Black group, the parallel analysis of random and real factors and the item factor loadings for alternative factor solutions also indicated that a three-factor solution was adequate.

Table 11
Eigenvalues of the Unrotated Random and Real Spatial Item Intercorrelation Matrices for Blacks and Whites, and Factor Contributions for the First Three Factors of the Rotated Factor Matrices

	Unrotated Solution				Rotated Solution		
			Percent of			Percent of	
Group and Factor	Random Data	Real Data	Common Variance	Total Variance	Factor Contribution	Common Variance	Total Variance
Blacks							
1	1.54	2.87	42.3	9.6	2.16	38.2	7.2
2	1.22	1.65	24.3	5.5	2.13	37.7	7.1
3	1.13	.123	18.1	4.1	1.36	24.1	4.5
4	1.08	1.03	15.2	3.4	----	----	---
Whites							
1	.99	3.57	49.7	11.9	2.40	43.6	8.0
2	.89	1.06	14.8	3.5	1.60	29.1	5.3
3	.80	.98	13.6	3.3	1.50	27.3	5.0
4	.70	.92	12.8	3.1	----	----	---
5	.65	.65	9.1	2.2	----	----	---

Although the number of factors needed for both groups was the same, the relative strength of the factors, as in the vocabulary test, differed for the two groups (see Table 11). After rotation, the first factor in the White group accounted for 8.0% of the total test variance in the White group and 7.2% in the Black group. However, in contrast to the factor solutions for the vocabulary items, the second factor for the Black group was stronger than that for the White group. Thus, for the spatial subtest, two relatively equal-strength factors were predominant in the Black group's data but not in the White group's, while for the vocabulary test, two relatively equal factors were obtained for the White group.

Table 12
Rotated Factor Loadings of the Spatial Test Items for Black and White Groups

Item	Blacks				Whites			
	Factor 1	Factor 2	Factor 3	h^2	Factor 1	Factor 2	Factor 3	h^2
1	.01	.10	-.23	.06	.13	.20	.30	.14
2	.20	.17	.07	.07	-.12	.51	-.08	.28
3	-.05	.45	-.20	.25	.18	-.07	.23	.09
4	.11	.27	.32	.19	.59	.02	.09	.36
5	.06	.02	.15	.03	.15	.18	.35	.19
6	.16	.09	.11	.05	.27	.00	.05	.07
7	-.01	.12	.11	.03	.07	.19	.34	.16
8	-.09	.53	-.06	.29	.22	.10	.18	.09
9	-.03	-.17	.19	.07	.03	.02	.18	.03
10	.40	.44	.10	.36	.06	.49	.24	.31
11	.14	.28	-.11	.10	.20	-.02	.38	.18
12	.33	-.03	-.04	.11	.20	.36	.27	.25
13	.12	.49	-.03	.26	.79	.13	-.17	.67
14	.09	.35	-.03	.13	.29	.17	.19	.15
15	.15	.40	-.02	.19	.15	-.01	.11	.03
16	-.39	.31	.21	.30	.23	.14	.26	.14
17	.10	.10	-.48	.25	.23	-.02	.43	.24
18	.19	.11	.26	.11	.34	.15	.12	.15
19	.18	.36	.21	.21	.11	.17	.30	.13
20	.62	.17	.04	.41	.32	.55	-.05	.41
21	.68	.22	-.02	.52	.55	.13	.12	.33
22	-.15	.07	.28	.11	.27	.13	.33	.21
23	-.04	-.09	.35	.13	-.03	-.22	-.09	.06
24	.76	.03	-.04	.58	.43	.52	-.06	.50
25	-.09	.32	.17	.14	.05	.16	.25	.09
26	.15	.32	.24	.18	.30	.19	.27	.20
27	.00	-.09	-.59	.35	.08	.11	.09	.03
28	.08	.24	-.04	.07	.01	-.02	.19	.04
29	.04	.22	.18	.08	.13	.05	.08	.03
30	-.17	-.15	.13	.07	-.04	.03	.14	.02
Factor Contribution	2.16	2.13	1.36	5.66	2.40	1.60	1.50	5.55

As indicated by the pattern of loadings in Table 12 and the factor correlations in Table 13, the second largest White factor seemed to define a factor

similar to the first Black spatial factor (r=.58), a result similar to that found in the vocabulary test. Beyond this factor pair, no clear correspondence existed between factors across the two groups, although the first White factor did share some items and variance with both the first and second largest Black factors; and the third White factor showed a small negative relationship to the first Black factor.

Table 13
Correlations of Rotated Factor
Loadings for Spatial Test Items

Factor in Black Group	Factor in White Group 1	2	3
1	.37*	.58***	-.33*
2	.33*	.04	.04
3	.01	-.01	-.10

*p<.05; **p<.01; ***p<.001

Summary. There were both similarities and differences between the two racial groups in the factor structures of the three subtests:

1. While the number of factors needed to account for the common variance in each test was the same for Blacks and Whites, the nature of these factors in terms of which items defined them or what they represented appeared to be different.

2. The largest factor for the White group accounted for more variance in all three tests than did the largest factor in the Black group. In addition, while the White group had two relatively large factors in the vocabulary data, the Black group had two relatively large factors in the spatial data.

3. While the factor structure of the mathematics test seemed to be similar for the two groups, the largest White factor in the vocabulary test, and to a lesser extent in the spatial test, seemed to have no comparably defined factor in the respective Black factor structures. In both the vocabulary and spatial tests, the largest factors in the Black group corresponded to the second largest factor in the White group.

Item Bias Analysis

Reckase (1977), drawing conclusions from a series of simulation studies, concluded that if the first unrotated factor in a set of test items accounts for at least 10% of the total variance and there are no strong competing unrotated factors, ICC item parameterization techniques will estimate the latent trait parameters indexing this first dimension. When the first unrotated factor accounts for a smaller percentage of the total test variance, however, the parameters will not be as stable as when the first factor accounts for a larger proportion of test variance. In the present vocabulary data, the first unrotated factors accounted for 15% and 10% of the total test

variance in the White and Black groups, respectively (see Table 5), and about 14% for the combined group. Therefore, based on Reckase's results, it can be assumed that the ICC parameters obtained from those items will estimate the first dimension of the vocabulary test items. Comparison of parameters across the Black and White groups, however, will have to take into consideration the possible instability of the estimates.

ICC item parameter values were estimated for 127 vocabulary test items. Table 14 shows the number of items failing the requirements of the parameterization procedure for each racial group. A total of 47 items did not fit the model for one or both of the groups. The results shown in Table 14 suggest (1) a greater tendency on the part of Black students not to engage in random guessing, as indicated by fewer items rejected with proportion correct less than .20, and (2) a tendency for the item responses of White students to correlate more highly with total scores on the test, as indicated by more items with very high point-biserial correlations with total score. This latter tendency would be related to the higher first factor loadings and the higher eigenvalue for the first factor observed for the White student group.

Table 14
Number of Items Rejected by the Parameterization Procedure for the Reasons Given, by Racial Group

Reason	Racial Group Whites	Blacks
Proportion Correct Less Than Fixed c Value of .20	14	28
Biserial Correlation Between Item and Item-Excluded Total Score Greater Than 1.0	13	5

For the remaining 80 items, there were four ICC parameters: (1) a discrimination parameter, a, for each racial group and the total group; (2) a difficulty parameter, b, for each racial group and the total group; (3) a guessing parameter, c, which was fixed for all items at .2 (1 divided by the number of multiple-choice response alternatives); and (4) an index of bias, derived by subtracting the White group's difficulty parameter from the Black group's difficulty parameter. A large positive value of the bias index, then, indicated an item which the Black group found to be more difficult on the average than the White, whereas a large negative bias index indicated an item which was more difficult for the White group. Four of the 80 items for which item parameters were obtained had extreme discrimination (a greater than 11.0) or extreme difficulty (b greater than 13.0) parameter values and were excluded from further analysis. Appendix Table A-5 contains the ICC parameters for the 80 vocabulary items. The last column of Table A-5 contains the index of bias for each item.

Table 15 shows the mean, standard deviation, and range of values of the discrimination (a) and difficulty (b) parameters for the Blacks, the Whites, and the combined group, as well as these statistics for the bias index. The

data in Table 15 show that although the items were, on the average, about equally discriminating for Blacks and Whites (mean a=.85 versus .87), they were more difficult for Blacks (mean b=.21 for Blacks versus -.03 for Whites). The higher average difficulty for the Black group was also reflected in the positive mean item bias (.35) which represents the average difference between the Black and White b-parameters. From the maximum and minimum bias values, it can be seen that the largest item difficulty difference in favor of the Whites was 4.01, whereas the largest difficulty difference in favor of Blacks was 5.46.

Table 15
Mean, Standard Deviation, and Range of the Discrimination (a), Difficulty (b), and Bias Parameters for Calibrated Items (N=76 items)

	Racial Group				Combined Racial Groups		
	Blacks		Whites				
Statistic	a	b	a	b	a	b	Bias
Mean	.85	.21	.87	-.03	.82	-.05	.35
Standard Deviation	.52	1.50	.50	1.70	.37	1.72	1.55
Maximum	2.87	3.61	3.06	6.83	2.27	3.98	4.01
Minimum	.13	-6.14	.08	-4.19	.09	-9.44	-5.46

Since the composition of the test included items written specifically for Blacks, it was relevant to determine the mean discriminations and mean difficulties for "Black-type" versus "non-Black-type" words separately. Table 16 reports these data as well as the mean bias for each set of words separately. Of the 76 items remaining after item calibration, 22 were "Black-type" words and the remainder of the item pool constituted the "non-Black-type" items. These data show that the "Black-type" words were less discriminating for both Blacks and Whites than the remainder of the items (mean a=.67 versus .92 for Blacks and .56 versus .99 for Whites). While the item pool as a whole was slightly more discriminating for Whites (see Table 15) and the "non-Black-type" words separately were slightly more discriminating for Whites (mean a=.99 for Whites versus .92 for Blacks), the "Black-type" words were more discriminating for Blacks (mean a=.67 for Blacks versus .56 for Whites).

Table 16
Mean, Standard Deviation, and Range of Discrimination (a), Difficulty (b), and Bias Parameters for "Black-Type" Words and "Non-Black-Type" Words

	Racial Group				Combined Racial Groups		
	Blacks		Whites				
Statistic	a	b	a	b	a	b	Bias
"Black-Type" Words (N=22)							
Mean	.67	-.23	.56	.52	.54	-.09	-.71
Standard Deviation	.51	1.39	.34	2.15	.36	2.56	2.08
Maximum	2.87	3.61	1.45	4.50	1.68	3.43	2.93
Minimum	.25	-2.25	.08	-4.19	.09	-9.44	-5.46
"Non-Black-Type" Words (N=54)							
Mean	.92	.37	.99	-.26	.93	-.03	.79
Standard Deviation	.49	1.54	.50	1.44	.32	1.23	.98
Maximum	2.75	3.17	3.06	6.83	2.27	3.98	4.01
Minimum	.13	-6.14	.14	-2.77	.22	-2.62	-4.44

As expected, the "non-Black-type" words were more difficult on the average for Blacks (mean *b*=.37 for Blacks versus -.26 for Whites), while the "Black-type" words were more difficult for Whites (mean *b*=.52 for Whites versus -.23 for Blacks). This was reflected further in the mean bias index which indicated that "Black-type" words favored Blacks (mean bias=-.71), while "non-Black-type" words favored Whites (mean bias=.79).

Relationship Between Bias Index and Factor Composition

In the discussion of comparative factor structures for Blacks and Whites, it was indicated that the largest Black factor in the vocabulary test corresponded most strongly to the second largest factor in the White group. The question was raised as to the nature of the largest White factor in the vocabulary test. Table 17, which shows the correlations between the bias index and each of the vocabulary factors in the Black and White groups, suggests one interpretation of why the largest White factor did not correspond to any Black factor. This table shows that none of the Black factors had item loadings which

Table 17
Correlations Between the Bias Index
and Item Factor Loadings for Each Vocabulary
Factor in Both Black and White Groups

Factor	Blacks	Whites
1	.20	.46**
2	-.05	.36**
3	-.08	.27*

*significant at $p \leq .05$
**significant at $p \leq .01$

correlated highly with the item indices of bias. On the other hand, all of the White factors were related to item bias to some degree, especially the first factor. Examination of the item loadings on this factor (compare Table 5 with Appendix Table A-5) showed that those items with high loadings on the first White factor were words drawn from the Minnesota vocabulary item pool and the ETS item pool. These items had bias values favoring Whites somewhat and ranged from about .50 to 2.00. The majority of the items with low or negative loadings on the first White factor were the "Black-type" words with negative bias indices.

This trend was also present to a lesser extent with the second White factor where some, but not all, "Black-type" words were associated with lower item loadings. Words written specifically for Whites did not seem to influence or define a factor in this way in either the White or Black groups.

The inclusion of the "Black-type" items seemed to have defined a "bias factor" in the White group, which was represented most strongly by the first factor. Not surprisingly, there was apparently not enough variation in performance on these items in the Black group to define such a factor. This interpretation is consistent with the strong relationship between the second White factor and the largest Black factor, and the relative independence of the largest White factor, which seemed to be largely a "bias factor."

Variance Overlap Analysis

Test score intercorrelations. Table 18 shows the correlations among scores on the three tests for the two racial groups and for the total combined group. For the Black group the correlation between vocabulary and spatial test scores was not significantly different from zero; for all other subtest combinations in both racial groups and the total combined group, subtest correlations were significantly different from zero. Test score correlations for the Black group were consistently lower than for the White group. However, statistical tests of the significance of these differences indicated that they were not significantly different from each other.

Table 18
Pearson Product-Moment Correlations Between Scores on the Vocabulary, Mathematics, and Spatial Tests

	Racial Group					Combined Racial Groups	
	Blacks		Whites				
Subtests	*N*	*r*	*N*	*r*	p^a	*N*	*r*
Vocabulary vs. Mathematics	48	.36*	96	.60*	.08	144	.53*
Vocabulary vs. Spatial	48	.16	96	.34*	.32	144	.31*
Mathematics vs. Spatial	81	.15*	180	.36*	.10	261	.37*

[a]Probability of error in rejecting null hypothesis of no difference in subgroup *r*'s.
*Significantly different from zero at $p<.05$.

One possible explanation for the consistently (but not significantly) lower test score intercorrelations for the Black group was suggested by the factor analyses, which indicated consistently lower proportions of variance extracted by the factor analytic procedure in the Black group. This suggests the possibility of either less reliable variance in the test scores of the Black group or additional reliable specific variance which was not being extracted by the factor analysis in the Black group. To investigate this possibility, Cronbach's alpha internal consistency reliability coefficients were calculated for each subtest total score in both groups.

Table 19
Alpha (α) Internal Consistency Reliabilities for Black and White Groups

	Blacks		Whites	
Test	*N*	α	*N*	α
Vocabulary	90	.89	169	.93
Mathematics	77	.59	179	.75
Spatial	83	.54	182	.71

Table 19 shows that the mathematics and spatial tests were less internally consistent in terms of Cronbach's alpha coefficient for Blacks than for Whites, while the internal consistency reliability coefficients of the vocabulary test

were about the same for both groups. This lower internal consistency could also explain to some degree the smaller observed score variances in test scores for the Black group, since the item covariances would contribute less to total test score variance for this group.

Factor score cross-correlations. The factor analyses of the three subtests reported above indicate that total scores on the subtests in each racial group were composed of two to three factors. Thus, the correlations of total raw scores combined several independent variables whose cross-correlations between tests might differ in the two racial groups. To explore this hypothesis, factor scores were computed in each group for the factors identified in each of the three tests. Factor scores were computed using the rotated factor loadings to weight each item's contribution to each factor. These factor scores were then correlated between tests separately for the two racial groups. The correlations are shown in Table 20.

Table 20
Cross-Correlations of Subtest Factor Scores
for Black (N=45) and White (N=94) Groups

	Vocabulary						Mathematics			
	Blacks			Whites			Blacks		Whites	
Factor	1	2	3	1	2	3	1	2	1	2
Mathematics										
1	.29*	.24*	.02	.21*	.34*	.29*				
2	-.05	.05	-.24*	.19*	.20*	.14*				
Spatial										
1	-.14	.15	-.09	.06	.34*	.19*	-.16	.10	.28*	.08
2	-.02	.05	-.16	.03	.27*	.01	-.10	-.03	.20*	-.02
3	.25*	.11	-.13	.34*	-.09	.13	.04	.34*	.31*	.22*

*p<.05

One major result is obvious from the data in Table 20: There was a tendency for factor scores in one subtest to correlate more highly with factor scores in the other subtests within the White group than within the Black group. For example, in comparing the vocabulary versus the mathematics test, there were a total of five factors involved for both groups (three vocabulary factors and two mathematics factors), allowing six possible pairings of vocabulary and mathematics factors. In the White group, five of the six factor pairings reflected shared variance between the two tests, whereas in the Black group only three of the six factor pairings were related. This trend was also evident in the other test pairs. In both the vocabulary versus spatial test, and mathematics versus spatial test comparisons, four factor pairings reflected significant relationships across the two tests for Whites, compared to only one factor pairing for Blacks. In addition to the larger number of significant correlations in the White group, the strength of these relationships between predictable factors across the tests was generally stronger in the White group.

The data in Table 20 support the results of the test score intercorrelation analysis as well as the earlier analysis of the correlations of factor

loadings between the two racial groups. The test score correlations showed that subtest scores were more highly correlated for Whites than for Blacks (see Table 18); the factor score intercorrelations described here support this result and suggest that the lower subtest intercorrelations for Blacks may in part be due to the smaller strength and number of predictable aspects (factors) of performance across the subtests. The correlations of factor loadings between racial groups (see Tables 7, 10, and 13) had suggested that different factors were being measured in the two groups with the same sets of items; the correlations of factor scores across subtests showed that the factors in different tests relate to each other differently within the two racial groups, thereby supporting the suggestion that different factors may be involved.

Redundancies. Although the intercorrelations among the test scores and factor scores for the two racial groups suggested higher interrelationships for the Whites than for the Blacks, the results in Tables 18 and 20 did not give a clear answer as to the amount of variance among subsets of scores in common between the tests in the two groups. To answer this question, a redundancy analysis (Weiss, 1972) was computed from the squared multiple correlations of the factor scores from each subtest predicting the factor scores in the other subtests within each racial group. Appendix Table A-6 gives the squared multiple correlations and the redundancy values for each pairing of subtests in each group; the redundancy data are summarized in Table 21.

Table 21
Proportions of Variance in a Set of Factor Scores Predicted by the Other Set of Factor Scores (Redundancies) for Black and White Groups

	Vocabulary vs. Mathematics		Vocabulary vs. Spatial		Mathematics vs. Spatial	
	Black	White	Black	White	Black	White
Redundancy of Second Test Factors on First Test Factors	7.0%	10.7%	5.3%	11.2%	8.0%	9.5%
Redundancy of First Test Factors on Second Test Factors	9.6%	16.8%	5.3%	13.1%	5.3%	6.6%

For both racial groups, the proportion of variance in the mathematics factors predicted by the vocabulary factors (versus the proportion of variance in the vocabulary factors predicted by the mathematics factors) indicated that the vocabulary factors were better predictors of the mathematics factors (redundancies of 9.6% and 16.8%) than vice versa. This is not surprising, since the majority of the mathematics items were of the word-problem variety and thus might be expected to be more influenced by verbal ability than the vocabulary test would be by mathematics ability.

The results for the two racial groups indicate a lower degree of common variance between the mathematics and vocabulary factors for the Blacks than for the Whites. In the Black group the vocabulary test factors accounted for 9.6% of the variance of the mathematics factors, while in the White group vocabulary accounted for 16.8% of the variance in the mathematics factor scores.

When vocabulary factors were predicted from mathematics factors, 7% of the variance was accounted for in the Black group, while 10.7% was accounted for in the White group.

A similar pattern of results was found for the other pairings of subtest factors. In general, the proportions of variance in a set of factor scores accounted for by the other test's factor scores was less in the Black group than in the White group. The largest differences occurred when the vocabulary test was one of the variable sets, particularly when it was the "predictor" set. This occurred in spite of the fact that the largest White factor, and to a lesser extent the second White factor, contained a "bias" component, i.e., correlated moderately with the item bias index, as shown above. From Table 20, however, it can be seen that the largest White vocabulary factor was not as predictive of the primary mathematics and spatial factors as the second and third White factors nor as predictive as the primary Black vocabulary factors in predicting the mathematics factors. This is consistent with the interpretation of the largest White vocabulary factor as being largely a "bias" factor, since such a factor would be expected to relate less to more legitimate factors in the other tests. Only small racial group differences in variance accounted for occurred in the relationship between the mathematics and spatial tests.

DISCUSSION AND CONCLUSIONS

This research was concerned with (1) the extent to which ability test performance of Blacks and Whites differed on a number of *kinds of ability test items*, (2) the analysis of some demographic variables related to observed differences, (3) the use of an index of item-by-group interaction to identify biased test items, (4) an analysis of racial group differences in the structure of their responses to ability test items, and (5) an analysis of racial group differences in relationships among performance on different ability domains.

Analysis of variance of total number-correct scores on the vocabulary, mathematics, and spatial subtests showed significant mean differences in scores for all subtests, with Whites averaging higher than Blacks. In the vocabulary test, however, this effect could not be interpreted independently of sex and parents' educational level. The race effect was consistent only for Black males who averaged lower than all other groups at all parental educational levels. The results for the three tests generally concurred with most other studies which have examined Black and White differences on mental ability tests. The results for the vocabulary test were most interesting, since more items were specifically oriented toward Blacks than toward Whites. The generally lower mean scores for Blacks on these three tests did not seem attributable to socioeconomic differences, at least as indexed by parents' educational level, since the groups were comparable on this variable.

In the vocabulary test, significant mean differences were found for parents' educational level and sex, but the significant two-way interaction of parents' educational level and sex and the race × education × sex three-way interaction reduced the interpretative significance of these differences. Number-correct scores generally increased with increasing parental education except for Black males. No sex or education effect was found in the mathematics test, while in the spatial test the main effect for parents' educational

level was statistically significant; but higher parental education was only associated with higher spatial scores in the White group. However, even for Whites, the size of the mean difference was small. Thus, parents' educational level was related more to vocabulary test scores than to mathematics or spatial scores.

While the race differences generally found in this study concur with most prior research, the results relating to the parental education and sex variables did not clearly replicate previous findings in all cases. Two recent books (Garai & Sheinfeld, 1968; MacCoby & Jacklin, 1974) have reviewed much of the literature on sex differences and concluded that the evidence supports male superiority on tests of mathematical and spatial abilities. While the two reviews drew somewhat differing conclusions on sex differences in various measured vocabulary abilities (e.g., analogies, word fluency, verbal reasoning, comprehension), both concluded that females have shown superior ability to males in word fluency, a type of vocabulary ability which would seem to be somewhat related to the vocabulary test used in the present study. The frequent finding of male superiority on tests of mathematical and spatial abilities reported by these authors was not replicated here. On the vocabulary test, Black females scores higher than Black males at all parental educational levels. However, for Whites there was an interaction of the sex variable with parental educational level, with females scoring higher at only the middle parental educational level. Thus, the usual finding of female dominance on this type of test was only partially replicated.

The strong relationship between measured intelligence and socioeconomic level is well documented (for a review, see Tyler, 1965). To the extent that the composite parental education index used in this study indexes socioeconomic class, higher intelligence levels would be expected in the students with higher parental educational levels due to a genetic contribution; and such differences would be reflected in the three tests used in this study.

In addition to the genetic contribution to higher test scores that would be predicted by higher parental intelligence levels, differences in the home environment provided by parents with varying degrees of education might be expected to be related to test scores, although some evidence suggests that this effect may contribute less to performance on the spatial test than to performance on the vocabulary and mathematics tests (Marjoribanks, 1972). Because of the small number of students reporting parents' education at various levels in this study, it was necessary to collapse the original six educational categories into three levels for the analyses performed. The resulting educational levels ("some high school," "high school diploma," "some post-high school ") did not provide as discriminating measures of parents' educational level as had been hoped for, and this might be expected to reduce any relationship between this variable and scores on the three ability tests.

Nonetheless, significant effects for the parental education variable were obtained in both the vocabulary and spatial tests, even though the mean differences in the spatial test were not of much practical significance. In addition, the trend of vocabulary scores over increasing educational levels depended on the race and sex group concerned. The education variable seemed important for all groups except the Black males.

No significant effect for parental education for the mathematics test was

found, contrary to what would be expected from previous research. The relatively low mean scores for all groups on this test suggest that this test might have been too difficult for the sample tested, which might have resulted in the lack of relationship with the parental education variable.

Application of ICC theory item parameterization techniques to the vocabulary data showed that item parameter estimates could be obtained for the majority of the items, although the first principal component represented only 15% of the total variance in the vocabulary test items in the White group and about 10% of the total vocabulary item variance in the Black group.

An item-by-group index of item bias was computed for each vocabulary item by subtracting the ICC item difficulty (b) parameter obtained in the White group from the difficulty parameter in the Black group. Results showed that, on the average, items were (1) more difficult for Blacks, (2) about equally discriminating in the two groups, and (3) more biased in favor of Whites.

Examination of the item parameter characteristics for "Black-type" words separately from the rest of the calibrated item pool indicated that the "Black-type" items were more discriminating for Blacks than for Whites, while the rest of the item pool was more discriminating for Whites. The "Black-type" words were less discriminating than the remainder of the items for both groups. Since average item discrimination is positively related to the magnitude of internal consistency reliability coefficients as well as to values of test information (both of which are indices which reflect the precision in responses to a set of test items), this result suggests that more precise estimates of the verbal ability of Blacks would *not* be obtained by using "Black-type" test items of the type used in this study rather than "non-Black-type" items.

A comparison of the factor structures for Whites and Blacks on the three subtests concluded that

1. There was substantial specific item variance in each of the three tests for both groups.

2. In all tests the strength of the first factor was larger for Whites than for Blacks.

3. While the number of important factors appeared to be the same in both groups, the patterns of item loadings on these factors indicated that the nature of the factors, except for the mathematics test, differed for the two groups. In particular, the largest vocabulary factor in the White group did not correspond to any similar factor in the Black group. This last result differs from most previous comparisons of factor structures in Black and White groups, which have generally found that dimensions measured were the same in the two groups but that the levels of performance on each dimension differed.

Many previous studies which found similar factor structures for Blacks and Whites utilized widely used standardized tests from which biased items have largely been eliminated based on previous analyses. In this study, items written specifically for Blacks and Whites were purposely included in the vocabulary test in order to test the methodology of identifying possibly biased items using an index of item bias derived from ICC theory and to investigate in a

subsequent study the use of computerized adaptive testing techniques to reduce this bias. While the items written specifically for Whites did not define a separate factor in either group's factor solutions, the largest White vocabulary factor seemed to be largely defined by the words written specifically for Blacks. No such "bias" factor appeared in the Black group, presumably because this group was more homogeneous with respect to these items. As a result, the finding of differential factor structures in the present study, in contrast to previous research, is not surprising. Indeed, the finding of differential factor structures when potentially biased items were introduced supports the utility of an index of item bias and comparative factor analyses in eliminating such differences. A subsequent study will evaluate the use of the bias index reported here in conjunction with computerized adaptive testing techniques to reduce differences in level and dimensions of performance between racial groups.

Three analyses were concerned with the question of the similarities between the two racial groups in variance overlap among the three kinds of ability tests. The raw scores on the three tests were intercorrelated for Blacks and Whites; all test intercorrelations were significant except for the correlation between vocabulary and spatial scores for Blacks. There was a consistent trend for the White correlations to be larger than the Black correlations, although the differences were not statistically significant.

In the second analysis concerned with variance overlap, factor scores computed separately for each racial group on each of the subtest factors were correlated with factor scores for each of the other subtests. These results supported both the correlations of total scores and the comparison of factor structures between the racial groups. The correlations between factor scores on different tests for the Blacks were consistently lower than for the Whites. In addition, there were different patterns of intercorrelations for the two racial groups, supporting other analyses which indicated that the same tests measure different components of ability in the two racial groups.

The redundancy analysis supported the other two variance overlap analyses, indicating that there was less shared variance between the subtests for the Black group than for the White group. The largest differences occurred for test pairings which included the vocabulary test, suggesting that differences in vocabulary ability are somehow important in reducing variance overlap in abilities for the Blacks. This finding is probably not due to the lower internal consistencies for the Blacks, since the difference in internal consistencies was smallest for the vocabulary test.

The data of this study indicate that using number-correct scores, Blacks generally performed less well on tests of vocabulary, mathematics, and spatial ability than did Whites; but additional data analysis indicated that a number of variables confounded the interpretation of this finding. In the vocabulary test, race, parental education, and sex all interacted with each other. Thus, to the extent that group mean differences on this test were in fact due to bias in the test, the exact nature and interpretation of this bias is complex. At the item level, more items in the vocabulary test were shown to be biased against Blacks than against Whites, as indexed by the difference in ICC difficulty parameters for the two groups. This occurred despite the inclusion of more words written specifically for Blacks than for Whites. With very few exceptions, the only words which favored Blacks (i.e., had a negative bias

index) were those written specifically for Blacks (see Appendix Table A-5). This suggests that either a majority of the items in the test were biased, or that the vocabulary ability level of the Black group was in fact lower than that of Whites in the sample.

Also confounding the interpretation of the racial group mean differences was the differential structure of two of the tests for the two racial groups. The data showed substantial group differences in both the within-test structures and the between-test relationships. It is clear from these data that the subtests did not measure the same factors in the two racial groups and that the organization of the abilities as measured by these subtests differed for the two groups. It was suggested that the largest White vocabulary factor was largely a "bias" factor, being defined in large part by the variance associated with White performance on "Black-type" items. Had these items not been included in the vocabulary test, the factor structures would likely have been more similar for the two groups.

The presence of group differences in factor structure for the tests as given, however, makes the meaningfulness of number-correct scores as a basis for group comparisons questionable. With the same test measuring different components of ability in the two racial groups, number-correct scores will be composed of these different abilities in different degrees. The effect, then, of directly comparing number-correct scores between the two groups is that of comparing "apples" and "oranges." A conclusion of meaningful racial group differences in mean ability scores can only be valid if it can be demonstrated that the tests are measuring the same variable for the two groups. Both the factor analyses and the variance overlap analysis indicate that the abilities measured here, especially the vocabulary and spatial abilities, were measured differently in Black and White students, even though the same tests were used; consequently, direct comparisons of these groups on number-correct scores is inappropriate. The use of an index of item bias in conjunction with factor and redundancy analyses can assist in determining whether number-correct scores represent the same components of ability for different groups and thus can be meaningfully compared.

REFERENCES

Angoff, W. H., & Ford, S. F. Item-race interaction on a test of scholastic aptitude. Princeton, NJ: Educational Testing Service, October 1971, pp. 1-24.

Bartlett, D. P., Newbrough, J. R., & Tulkin, S. R. Raven Progressive Matrices: An item and set analysis of subjects by race, sex, and social class. Journal of Consulting and Clinical Psychology, 1972, 38, 154.

Baughman, E. E., & Dahlstrom, W. G. Negro and white children: A psychological study in the rural south. New York: Academic Press, 1968.

Breland, H. M., Stocking, M., Pinchak, B. M., & Abrams, N. The cross-cultural stability of mental test items: An investigation of response patterns for ten socio-cultural groups (Research Report PR-74-2). Princeton, NJ: Educational Testing Service, February 1974.

Cardall, D., & Coffman, W. E. A method for comparing the performance of different groups on the items in a test (Research Bulletin 64-61). Princeton, NJ: Educational Testing Service, 1964.

Cleary, T. A., & Hilton, T. L. An investigation of item bias. Educational and Psychological Measurement, 1968, 28, 61-75.

Flaugher, R. L., & Rick, D. A. Patterns of ability factors among four ethnic groups. Proceedings of the 80th Annual Convention of the American Psychological Association, 1972, 7, 27-28.

Garai, J. E., & Scheinfield, A. Sex differences in mental and behavioral traits. Genetic Psychology Monographs, 1968, 77, 169-299.

Gitlitz, A. H., & Kaufman, N. L. Influence of race, sex, and city on inductive reasoning items (Project Access Research Report No. 4). Princeton, NJ: Educational Testing Service, June 1972, pp. 1-23.

Humphreys, L. G., & Ilgen, D. R. Note on a criterion for the number of common factors. Educational and Psychological Measurement, 1969, 29, 571-578.

Humphreys, L. G., & Taber, T. Ability factors as a function of advantaged and disadvantaged groups. Journal of Educational Measurement, 1973, 10, 107-115.

Jensema, C. A simple technique for estimating latent trait mental test parameters. Educational and Psychological Measurement, 1976, 36, 705-716.

Jensen, A. R. Educability and group differences. New York: Harper & Row, 1973.

Jensen, A. R. How biased are culture-loaded tests? Genetic Psychology Monographs, 1974, 90, 185-244.

Johnson, M. Factorial invariance of African educational abilities and aptitudes (Research Bulletin RB-69-3). Princeton, NJ: Educational Testing Service, 1969.

Kennedy, W. A., Van de Riet, V., & White, J. C. A normative sample of intelligence and achievement of Negro elementary school children in the southeastern United States, Monographs of the Society for Research in Child Development, 1963, 28, No. 6.

Lesser, G. S., Fifer, G., & Clark, D. H. Mental abilities of children from different social-class and cultural groups. Monographs of the Society for Research in Child Development, 1965, 30, No. 4.

Loehlin, J. C., Lindzey, G., & Spuhler, J. N. Race differences in intelligence. San Francisco: Freeman, 1975.

Lord, F. M., & Novick, M. R. Statistical theories of mental test scores. Reading, MA: Addison-Wesley, 1968.

Lumsden, J. Test theory. Annual Review of Psychology, 1976, 22, 254-280.

MacCoby, E.E., & Jacklin, C. N. The psychology of sex differences. Stanford, CA: Stanford University Press, 1974.

Marjoribanks, K. Environment, social class, and mental abilities. Journal of Educational Psychology, 1972, 63, 103-109.

McBride, J.R., & Weiss, D. J. A word knowledge item pool for adaptive ability measurement (Research Report 74-2). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, June 1974. (NTIS No. AD 781894).

McDonald, R. P., & Ahlawat, K. S. Difficulty factors in binary data. British Journal of Mathematical and Statistical Psychology, 1974, 27, 82-99.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. Statistical package for the social sciences (2nd ed.). New York: McGraw Hill, 1975.

Reckase, M. D. Ability estimation and item calibration using the one- and three-parameter logistic models: A comparative study (Research Report 77-1). Columbia, MO: University of Missouri, Educational Psychology Department, Tailored Testing Research Laboratory, November 1977.

Shuey, A. M. The testing of Negro intelligence. Lynchburg, VA: J. P. Bell, 1958.

Stewart, D., & Love, W. A general canonical correlation index. Psychological Bulletin, 1968, 70, 160-163.

Stodolsky, S. S., & Lesser, G. Learning patterns in the disadvantaged. Harvard Educational Review, 1967, 37, 546-593.

Terman, L. M., & Merrill, M. A. Revised Stanford-Binet Intelligence Scale: Third Edition. Boston: Houghton Mifflin, 1960.

Tyler, L. E. The psychology of individual differences (3rd edition). Appleton-Century-Crofts, Inc., New York, 1965.

Urry, V. W. Ancillary estimators for the item parameters of mental test models. Washington, DC: U.S. Civil Service Commission, Personnel Research and Development Center, 1974.

Weiss, D. J. Canonical correlation analysis in counseling psychology research. Journal of Counseling Psychology, 1972, 19, 241-252.

APPENDIX A: SUPPLEMENTARY TABLES

Table A-1
Sources for Black Vocabulary Words

Giovanni, Nikki.	Black Judgement. Detroit, Michigan: Broadside Press, 1968.
Hughes, Langston.	Selected Poems. New York: Knopf, 1966.
Hughes, Langston.	Good Morning Revolution. New York: Lawrence Hill, 1973.
Major, Clarence.	Dictionary of Afro-American Slang. New York: International Publishers, 1971.
Ellison, Ralph.	Invisible Man. New York: Signet, 1947.
Little, Malcolm.	Malcolm X on Afro-American History. New York: Pathfinder Press, 1970.
Little, Malcolm.	By Any Means Necessary. (Edited by George Breitman). New York: Pathfinder Press, 1970).

Table A-2
Examples of Vocabulary Items

Items from Black Literature and Black Psychologist

Ranking	Gatemouth
1. Murdering	1. Gossiper
2. Exchange of insults	2. Doorway
3. Pig's intestines	3. Jazz musician
4. Fried cow's tail	4. Dog
5. Olympic event	5. Fat person

Shiv	Swag
1. Politician	1. Construction worker
2. Genius	2. Beggar
3. Book	3. Corrupt politician
4. Drifter	4. Stolen goods
5. Knife	5. Garbage

"White" type items from Webster's Seventh Collegiate Dictionary

Borsch	Torte
1. Overcoat	1. Cake
2. Dog	2. Twist
3. Porter	3. Shirt
4. Soup	4. Crime
5. Chamber	5. Answer

Afghan	Gefilte Fish
1. Alien	1. Type of fish
2. Harbor	2. A game
3. Canvas	3. Food
4. Vista	4. A sport
5. Blanket	5. Sucker

Items from Standardized Vocabulary Tests

Accumulate	Reinforce
1. Become cloudy	1. Speak loudly
2. Get angry	2. Come again to
3. Get dirty	3. Revise
4. Imitations	4. Apply again
5. Claws	5. Make stronger

Oppressed	Capitulate
1. Wrinkled	1. Entitle
2. Expressed	2. Surrender
3. Musically talented	3. Behead
4. Disowned	4. Put in charge
5. Put down	5. Congratulate

Table A-3
Sample Items from the Mathematics Test

Nancy bought 8 feet of ribbon. The ribbon was taken from a roll 4 yards long. How many inches were left on the roll after the sale?
A. 4 inches
B. 24 inches
C. 36 inches
D. 48 inches
E. none of these

John bought a bicycle for 1/5 as much money as Sam. Sam paid $85.00. How much did John pay?
A. $17.00
B. $40.15
C. $68.00
D. $80.00
E. none of these

The milk man had two containers. One was filled with 3.3 quarts of milk, and the other with 2.199 quarts. How much more milk was in one container than the other?
A. 1.101 quarts
B. 1.201 quarts
C. 2.101 quarts
D. 5.499 quarts
E. none of these

The price of a new car is $4565.08. Las year the car was $577.79 less. How much was last year's model?
A. $2989.29
B. $3087.29
C. $3887.29
D. $3987.29
E. none of these

A tent is 18 feet long and 14 feet wide. How many square feet does the floor of the tent cover?
A. 32 square feet
B. 64 square feet
C. 152 square feet
D. 252 square feet
E. none of these

A roll of paper 8 1/4 feet wide is to be cut into strips 9 inches wide. How many strips can be cut from this roll?
A. 9 strips
B. 10 2/3 strips
C. 11 strips
D. 17 1/4 strips
E. none of these

A television normally sells for $596. During a sale it was reduced by 22%. By how much was the item reduced?
A. $119.20
B. $121.12
C. $131.12
D. $464.88
E. none of these

One quart of lemonade was poured from a 7 gallon jug. How many pints of lemonade are left in the jug?
A. 48 pints
B. 54 pints
C. 64 pints
D. 108 pints
E. none of these

Table A-4
Percentages of Highest Educational Level Attained by
Fathers and Mothers as Reported by Students in Each Test Group

Highest Educational Level Attained	Racial Group: Blacks Father	Racial Group: Blacks Mother	Racial Group: Whites Father	Racial Group: Whites Mother	Combined Racial Groups Father	Combined Racial Groups Mother
Vocabulary Test Group	(*N*=90)		(*N*=167)		(*N*=257)	
Not Reported	14.45	8.9	3.0	1.8	7.0	4.3
Responded "Not Applicable"	15.6	10.0	7.2	2.4	10.1	5.1
Did Not Complete High School	16.7	15.6	28.1	18.6	24.1	17.5
Finished High School	25.6	25.6	35.3	55.1	31.9	44.7
Some College or Post-High School	16.7	22.2	13.2	11.4	14.4	15.2
Finished College	5.6	11.1	9.6	8.4	8.2	9.3
Attended Graduate or Professional School	5.6	6.7	3.6	2.4	4.3	3.9
Math and Spatial Test Group	(*N*=81)		(*N*=180)		(*N*=261)	
Not Reported	19.8	24.7	7.8	5.6	11.5	11.5
Responded "Not Applicable"	18.5	11.1	7.8	5.0	11.1	6.9
Did Not Complete High School	17.3	12.3	25.0	17.2	22.6	15.7
Finished High School	19.7	23.5	40.6	56.1	34.1	46.0
Some College or Other Post High School	16.0	19.8	7.8	10.0	10.3	13.0
Finished College	8.6	3.7	6.7	5.6	7.3	5.0
Attended Graduate or Professional School	0.0	4.9	4.4	.6	3.1	1.9

Table A-5
Item Characteristic Curve Discrimination (a) and Difficulty (b) Parameters for the Black, White, and Combined Groups, and the Corresponding Bias Index for the Total Combined Group

Item Stem	Item Number	Blacks (N=92)		Whites (N=173)		Combined Group (N=265)		Bias
		a	b	a	b	a	b	
1. Torte	1201	1.24	1.27	.85	.93	.98	1.01	.33
2. Ranking	1202	.30	.58	.34	.68	.33	.65	-.11
3. Pulpit	1204	.25	.82	.50	.41	.43	.52	.41
4. Sober	1205	.74	-1.95	.56	-2.71	.65	-2.31	.77
5. Shiv	1206	.43	1.51	.18	1.41	.28	1.54	.09
6. Evict	1207	1.23	.43	1.26	.00	1.34	.16	.44
7. Turmoil	1208	.34	.94	1.07	.02	.74	.22	.92
8. Discuss	1209	.69	-2.50	.71	-2.59	.67	2.67	.09
9. Contain	1210	.13	-6.14	.39	-2.13	.32	-2.62	4.01
10. Wail	1211	2.75	.36	2.17	.07	2.27	.18	.29
11. Triumph	1212	1.27	-.11	.87	-1.44	.95	-.87	1.33
12. Backgammon	1213	.53	1.37	1.17	.24	.98	.47	1.12
13. Igg	1214	.00	59.27	.16	3.53	.04	8.48	55.74
14. Indictment	1219	.64	1.31	1.31	.28	1.10	.38	.41
15. Shouting	1223	.45	-1.23	.08	4.23	.09	.84	-5.46
16. Unemployed	1224	.71	-2.29	.87	-2.77	.79	-2.55	.48
17. Sanctuary	1225	.76	.73	.78	.33	.84	.48	.40
18. Corrupt	1227	.55	1.07	1.77	.17	1.19	.34	.90
19. Oppressed	1228	.85	.36	1.19	.17	1.07	.26	.19
20. Diplomatic	1231	1.80	.73	.69	-.04	.85	.25	.77
21. Fry	1232	.59	.87	.24	4.23	.26	3.17	-3.36
22. Long Money	1233	.06	13.86	.29	3.13	.19	4.48	10.73
23. Lobster	1234	1.29	-1.47	.76	-2.13	.90	-1.83	.66
24. Taters	1235	.88	-.94	.75	-.95	.79	-9.44	.01
25. Enemy	1236	.49	.09	.89	-.42	.75	-2.55	.51
26. Cameo	1238	.38	3.17	.89	.20	.79	.64	2.97
27. Yam	1239	.90	-1.85	.49	-1.49	.52	-1.66	-.36
28. Exploit	1240	.82	1.47	.97	.72	.97	.91	.74
29. Chitterlings	1242	.84	-2.06	.28	2.51	.09	.81	-4.57
30. Ascot	1243	.56	2.30	1.53	1.02	1.22	1.23	1.28
31. Domination	1244	.61	.78	1.43	.27	1.06	.39	.51
32. Clip	1245	.32	2.39	.14	6.83	.22	3.98	-4.44
33. Ballerina	1246	1.44	-.56	.74	-1.19	.93	-.88	.62
34. Capitulate	1247	.07	19.93	.10	16.19	.10	16.84	3.74
35. Slide	1248	.25	-1.34	.41	-.23	.35	-.50	-1.11
36. Scratch	1249	.55	3.61	.60	2.08	.64	2.31	1.53
37. Rigid	1251	.66	2.49	1.31	.48	1.11	.84	2.01
38. Uppity	1252	.71	-.44	.54	.10	.56	-.12	-.54
39. Linoleum	1254	2.87	-.90	1.45	-1.05	1.68	-.60	.95
40. Macking	1257	.70	-1.24	.62	1.84	.22	1.61	-3.08
41. Badger	1259	.66	-.71	2.18	-1.01	1.20	-.91	.30
42. Lynch	1262	.70	-1.16	.86	-.81	.76	-.94	-.35
43. Hew	1264	.69	3.09	.47	2.57	.58	2.54	.52
44. Roust	1265	.59	1.53	1.18	-.03	1.03	.32	1.55
45. Gig	1268	.54	.08	1.01	-1.85	-.03	.00	.11

(continued on next page)

Table A-5(continued)
Item Characteristic Curve Discrimination (*a*) and Difficulty (*b*) Parameters for the Black, White, and Combined Groups, and the Corresponding Bias Index for the Total Combined Group

Item Stem	Item Number	Blacks (*N*=92)		Whites (*N*=173)		Combined Group *N*=265)		Bias
		a	*b*	*a*	*b*	*a*	*b*	
46. Testify	1272	.66	-1.54	.55	-1.85	.63	-1.59	.31
47. Rap	1275	.60	-2.25	.51	-2.21	.56	-2.04	-.03
48. Cracklins	1278	.58	.53	.24	4.50	.22	3.43	-3.97
49. Agile	1279	.40	2.34	.81	.73	.68	1.05	1.61
50. Gatemouth	1280	.49	.46	.96	.29	.75	.34	.16
51. Aardvark	1281	.94	.44	.74	-.96	.84	-.38	1.40
52. Rehearse	1282	1.47	-.47	.93	-1.27	1.08	-.91	.79
53. Custody	1284	.88	.51	1.06	.44	1.01	-.09	.94
54. Lush	1286	11.22	1.12	.53	3.59	.77	2.29	-2.47
55. Retrieve	1287	.84	.11	.61	-.94	.73	-.47	1.05
56. Imply	1289	.53	1.27	1.28	-.11	1.07	.17	1.39
57. Irritate	1293	1.78	-.14	.80	-.78	1.07	-.47	.64
58. Gay	1294	.57	-1.26	.27	-4.19	.39	-2.43	2.93
59. Dune	1295	.90	-.13	.89	-1.22	.91	-.77	1.08
60. Analyze	1297	.71	1.18	.74	.14	.76	.45	1.04
61. Solitary	1298	1.26	.09	.86	-.51	1.01	-.28	.60
62. Croak	1299	.34	-.18	.66	.02	.53	-.03	-.20
63. Blunder	1300	.98	1.18	1.12	.54	1.22	.71	.64
64. Pig	1301	.37	-.54	.28	-.33	.28	-.53	-.21
65. Bequeath	1302	.63	1.84	3.06	.64	1.49	.87	1.21
66. Afghan	1304	.87	.20	.95	-.57	.94	-.31	.77
67. Arbitrate	1305	1.02	1.91	1.83	.88	1.33	1.14	1.03
68. African dominoes	1309	.72	.95	.34	1.25	.45	1.10	-.30
69. Allow	1310	1.17	-.61	.84	-.95	.91	-.85	.35
70. Triple	1311	.92	-.58	.62	-1.93	.77	-1.22	1.35
71. Symbol	1312	.72	-.49	.51	-1.31	.62	-.91	.82
72. Reinforce	1313	.82	.06	.83	-.41	.89	-.23	.47
73. Specify	1315	1.32	.55	1.02	-.39	1.16	-.05	.95
74. Clarify	1316	.42	1.33	.67	-.13	.65	.25	1.46
75. Gadget	1320	1.21	-.24	1.41	-.84	1.31	-.62	.59
76. Caravan	1321	2.23	1.02	.80	.49	.96	.68	.52
77. Minimum	1322	.99	-.46	.57	-1.48	.74	-.96	1.02
78. Decode	1323	1.39	.01	1.14	-.68	1.19	-.40	.68
79. Tense	1324	1.32	.29	.61	-.47	.84	-.10	.76
80. Precede	1325	.82	1.47	.81	.38	.87	.69	1.09

Table A-6
Squared Multiple Correlations of Factor Scores in Each Subtest Predicted from Factor Scores in Other Subtests, and Redundancy Values, by Racial Group

Variables	R^2 and Redundancies: Blacks	R^2 and Redundancies: Whites
Prediction of Vocabulary (V) from Mathematics (M) Factor Scores		
V1 from M1, M2	.089	.068
V2 from M1, M2	.062	.139
V3 from M1, M2	.159	.095
Redundancy	7.0%	10.7%
Prediction of Mathematics (M) from Vocabulary (V) Factor Scores		
M1 from V1, V2, V3	.131	.242
M2 from V1, V2, V3	.062	.093
Redundancy	9.6%	16.8%
Prediction of Vocabulary (V) from Spatial (S) Factor Scores		
V1 from S1, S2, S3	.09	.118
V2 from S1, S2, S3	.03	.159
V3 from S1, S2, S3	.04	.058
Redundancy	5.3%	11.2%
Prediction of Spatial (S) from Vocabulary (V) Factor Scores		
S1 from V1, V2, V3	.053	.168
S2 from V1, V2, V3	.026	.072
S3 from V1, V2, V3	.087	.154
Redundancy	5.3%	13.1%
Prediction of Mathematics (M) from Spatial (S) Factor Scores		
M1 from S1, S2, S3	.034	.127
M2 from S1, S2, S3	.126	.059
Redundancy	8.0%	9.5%
Prediction of Spatial (S) from Mathematics (M) Factor Scores		
S1 from M1, M2	.036	.081
S2 from M1, M2	.012	.041
S3 from M1, M2	.118	.081
Redundancy	5.3%	6.6%

APPENDIX B: SUPPLEMENTARY FIGURES

Figure B-1
Sample Items from the Spatial Test

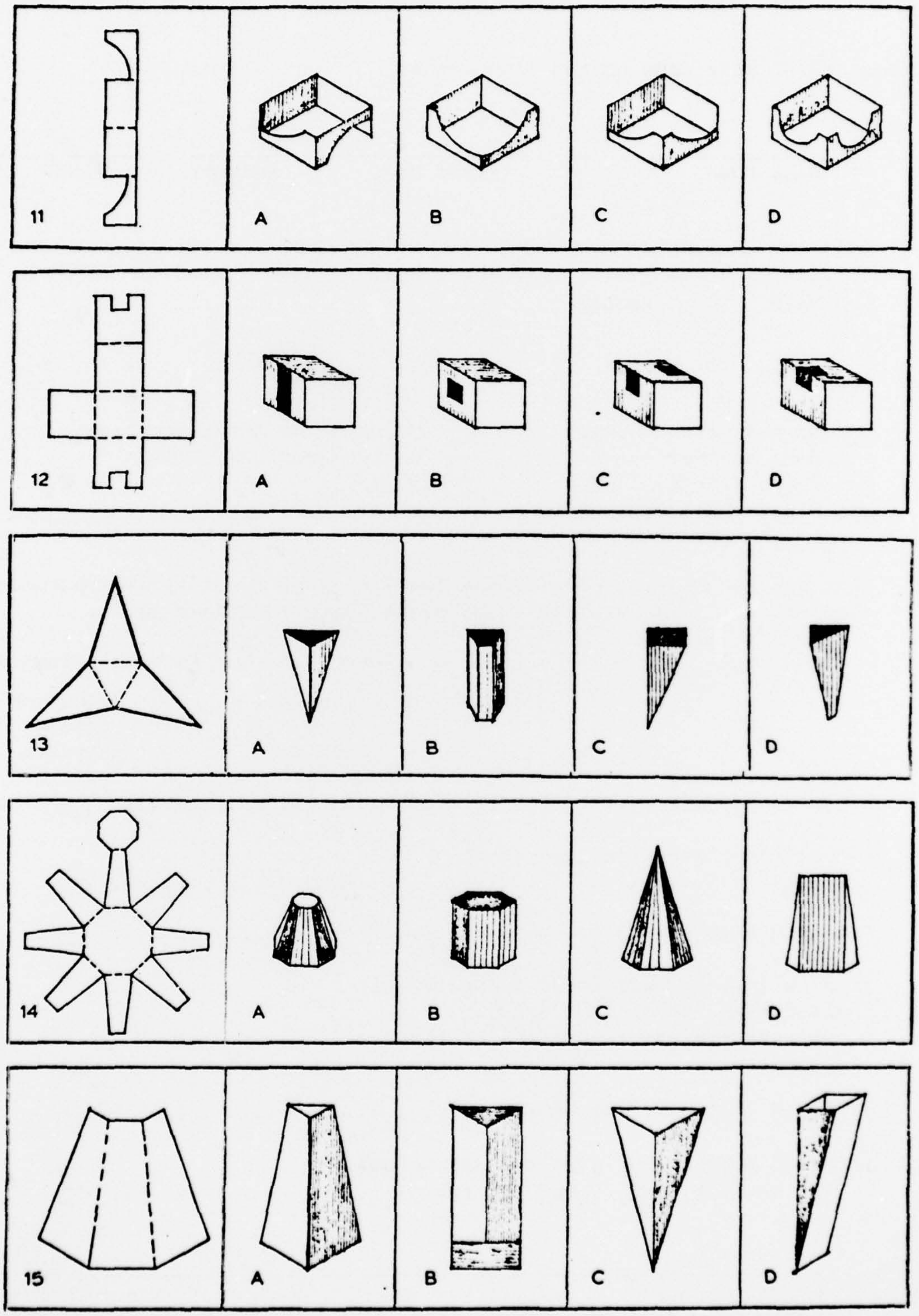


Figure B-2
Biographical Questionnaire

FOR RESEARCH PURPOSES ONLY

Please PRINT your name on the line below:

Last Name	First Name	Initial	Class

1. Sex

☐ Male ☐ Female

2. Race

☐ Afro-American (Black)
☐ Mexican-American
☐ Puerto-Rican
☐ Other Latin-American
☐ Oriental or Asian-American
☐ Native-American (Indian)
☐ White
☐ Other, specify: ____________

3. What was the *highest* educational level each of the following persons completed? If you are not sure, please give your best guess.

(Check one box in each column.)

	Father or male guardian	Mother or female guardian
Doesn't apply	☐	☐
Did not complete high (secondary school	☐	☐
Finished high school or equivalent	☐	☐
Some college or additional schooling after High School (for example, computer training, business or trade schools)	☐	☐
Finished college (four years)	☐	☐
Attended graduate or professional school (for example, law or medical school)	☐	☐

DO NOT TURN THE PAGE UNTIL YOU ARE TOLD TO DO SO.

DISTRIBUTION LIST

Navy

1 DR. JACK ADAMS
OFFICE OF NAVAL RESEARCH BRANCH
223 OLD MARYLEBONE ROAD
LONDON, NW, 15TH ENGLAND

1 Dr. Ed Aiken
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Jack R. Borsting
Provost & Academic Dean
U.S. Naval Postgraduate School
Monterey, CA 93940

1 Dr. Robert Breaux
Code N-71
NAVTRAEQUIPCEN
Orlando, FL 32813

1 DR. MAURICE CALLAHAN
NODAC (CODE 2)
DEPT. OF THE NAVY
BLDG. 2, WASHINGTON NAVY YARD
(ANACOSTIA)
WASHINGTON, DC 20374

1 Dept. of the Navy
CHNAVMAT (NMAT 034D)
Washington, DC 20350

1 Chief of Naval Education and
Training Support)-(01A)
Pensacola, FL 32509

1 Dr. Charles E. Davis
ONR Branch Office
536 S. Clark Street
Chicago, IL 60605

1 Dr. Richard Elster
Naval Postgraduate School
Monterrey, CA 93940

5 Dr. Marshall J. Farr, Director
Personnel & Training Research Programs
Office of Naval Research (Code 458)
Arlington, VA 22217

1 DR. PAT FEDERICO
NAVY PERSONNEL R&D CENTER
SAN DIEGO, CA 92152

1 CDR John Ferguson, MSC, USN
Naval Medical R&D Command (Code 44)
National Naval Medical Center
Bethesda, MD 20014

1 Dr. John Ford
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Eugene E. Gloye
ONR Branch Office
1030 East Green Street
Pasadena, CA 91101

1 CAPT. D.M. GRAGG, MC, USN
HEAD, SECTION ON MEDICAL EDUCATION
UNIFORMED SERVICES UNIV. OF THE
HEALTH SCIENCES
6917 ARLINGTON ROAD
BETHESDA, MD 20014

1 Dr. Steve Harris
Code L522
NAMRL
Pensacola FL 32508

1 CDR Robert S. Kennedy
Naval Aerospace Medical and
Research Lab
Box 29407
New Orleans, LA 70189

1 Dr. Norman J. Kerr
Chief of Naval Technical Training
Naval Air Station Memphis (75)
Millington, TN 38054

1 Dr. Leonard Kroeker
Navy Personnel R&D Center
San Diego, CA 92152

1 CHAIRMAN, LEADERSHIP & LAW DEPT.
DIV. OF PROFESSIONAL DEVELOPMMENT
U.S. NAVAL ACADEMYY
ANNAPOLIS, MD 21402

1 Dr. James Lester
ONR Branch Office
495 Summer Street
Boston, MA 02210

1 Dr. William L. Maloy
Principal Civilian Advisor for
Education and Training
Naval Training Command, Code 00A
Pensacola, FL 32508

1 Dr. James McBride
Code 301
Navy Personnel R&D Center
San Diego, CA 92152

2 Dr. James McGrath
Navy Personnel R&D Center
Code 306
San Diego, CA 92152

1 Commanding Officer
U.S. Naval Amphibious School
Coronado, CA 92155

1 Commanding Officer
Naval Health Research
Center
Attn: Library
San Diego, CA 92152

1 CDR PAUL NELSON
NAVAL MEDICAL R& D COMMAND
CODE 44
NATIONAL NAVAL MEDICAL CENTER
BETHESDA, MD 20014

1 DR. RICHARD J. NIEHAUS
CODE 301
OFFICE OF CIVILIAN PERSONNEL
NAVY DEPT
WASHINGTON, DC 20390

1 Library
Navy Personnel R&D Center
San Diego, CA 92152

6 Commanding Officer
Naval Research Laboratory
Code 2627
Washington, DC 20390

1 OFFICE OF CIVILIAN PERSONNEL
(CODE 26)
DEPT. OF THE NAVY
WASHINGTON, DC 20390

1 JOHN OLSEN
CHIEF OF NAVAL EDUCATION &
TRAINING SUPPORT
PENSACOLA, FL 32509

1 Office of Naval Research
Code 200
Arlington, VA 22217

1 Office of Naval Research
Code 437
800 N. Quincy SStreet
Arlington, VA 22217

1 Scientific Director
Office of Naval Research
Scientific Liaison Group/Tokyo
American Embassy
APO San Francisco, CA 96503

1 Dr. Kneale Marshall
SCIENTIFIC ADVISOR TO THE CHIEF
OF NAVAL PERSONNEL
NAVAL BUREAU OF PERSONNEL (PERS OR)
RM. 4410, ARLINGTON ANNEX
WASHINGTON, DC 20370

1 DR. RICHARD A. POLLAK
ACADEMIC COMPUTING CENTER
U.S. NAVAL ACADEMY
ANNAPOLIS, MD 21402

1 Mr. Arnold I. Rubinstein
Human Resoureces Program Manager
Naval Material Command (0344)
Room 1044, Crystal Plaza #5
Washington, DC 20360

1 Dr. Worth Scanland
Chief of Naval Education and Training
Code N-5
NAS, Pensacola, FL 32508

1 A. A. SJOHOLM
TECH. SUPPORT, CODE 201
NAVY PERSONNEL R& D CENTER
SAN DIEGO, CA 92152

1 Mr. Robert Smith
Office of Chief of Naval Operations
OP-987E
Washington, DC 20350

1 Dr. Alfred F. Smode
Training Analysis & Evaluation Group
(TAEG)
Dept. of the Navy
Orlando, FL 32813

1 CDR Charles J. Theisen, JR. MSC, USN
Head Human Factors Engineering Div.
Naval Air Development Center
Warminster, PA 18974

1 W. Gary Thomson
Naval Ocean Systems Center
Code 7132
San Diego, CA 92152

1 DR. H.M. WEST III
DEPUTY ADCNO FOR CIVILIAN PLANNING
AND PROGRAMMING
RM. 2625, ARLINGTON ANNEX
WASHINGTON, DC 20370

1 DR. MARTIN F. WISKOFF
NAVY PERSONNEL R& D CENTER
SAN DIEGO, CA 92152

Army

1 ARI Field Unit-Leavenworth
P.O. Box 3122
Ft. Leavenworth, KS 66027

1 HQ USAREUE & 7th Army
ODCSOPS
USAAREUE Director of GED
APO New York 09403

1 DR. RALPH CANTER
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 DR. RALPH DUSEK
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Dr. Milton S. Katz
Individual Training & Skill
Evaluation Technical Area
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Harold F. O'Neil, Jr.
ATTN: PERI-OK
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 DR. JAMES L. RANEY
U.S. ARMY RESEARCH INSTITUTE
5001 EISENHOWER AVENUE
ALEXANDRIA, VA 22333

1 Director, Training Development
U.S. Army Administration Center
ATTN: Dr. Sherrill
Ft. Benjamin Harrison, IN 46218

1 Dr. Joseph Ward
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Air Force

1 Air Force Human Resources Lab
AFHRL/PED
Brooks AFB, TX 78235

1 Air University Library
AUL/LSE 76/443
Maxwell AFB, AL 36112

1 CDR. MERCER
CNET LIAISON OFFICER
AFHRL/FLYING TRAINING DIV.
WILLIAMS AFB, AZ 85224

1 Personnel Analysis Division
HQ USAF/DPXXA
Washington, DC 20330

1 Dr. Marty Rockway (AFHRL/TT)
Lowry AFB
Colorado 80230

1 Major Wayne S. Sellman
Chief, Personnel Testing
AFMPC/DPMYPT
Randolph AFB, TX 78148

1 Jack A. Thorpe, Capt, USAF
Program Manager
Life Sciences Directorate
AFOSR
Bolling AFB, DC 20332

1 Brian K. Waters, LCOL, USAF
Air University
Maxwell AFB
Montgomery, AL 36112

Marines

1 Director, Office of Manpower Utilization
HQ, Marine Corps (MPU)
BCB, Bldg. 2009
Quantico, VA 22134

1 DR. A.L. SLAFKOSKY
SCIENTIFIC ADVISOR (CODE RD-1)
HQ, U.S. MARINE CORPS
WASHINGTON, DC 20380

CoastGuard

1 MR. JOSEPH J. COWAN, CHIEF
PSYCHOLOGICAL RESEARCH (G-P-1/62)
U.S. COAST GUARD HQ
WASHINGTON, DC 20590

Other DoD

1 Dr. Stephen Andriole
ADVANCED RESEARCH PROJECTS AGENCY
1400 WILSON BLVD.
ARLINGTON, VA 22209

12 Defense Documentation Center
Cameron Station, Bldg. 5
Alexandria, VA 22314
Attn: TC

1 Dr. Dexter Fletcher
ADVANCED RESEARCH PROJECTS AGENCY
1400 WILSON BLVD.
ARLINGTON, VA 22209

1 Military Assistant for Human Resource
Office of the Director of Defense
Research & Engineering
Room 3D129, the Pentagon
Washington, DC 20301

1 Director, Research & Data
OSD/MRA&L (Rm. 3B919)
The Pentagon
Washington, DC 20301

1 Mr. Fredrick W. Suffa
MPP (A&R)
2B269
Pentagon
Washington, D.C. 20301

Civil Govt

1 Dr. Susan Chipman
Basic Skills Program
National Institute of Education
1200 19th Street NW
Washington, DC 20208

1 Dr. Lorraine D. Eyde
Personnel R&D Center
U.S. Civil Service Commission
1900 EStreet NW
Washington, D.C. 20415

1 Dr. William Gorham, Director
Personnel R&D Center
U.S. Civil Service Commission
1900 E Street NW
Washington, DC 20415

1 Dr. Joseph Markowitz
Office of Research and Development
Central Intelligence Agency
Washington, DC 20205

1 Dr. Andrew R. Molnar
Science Education Dev.
and Research
National Science Foundation
Washington, DC 20550

1 Dr. Vern W. Urry
Personnel R&D Center
U.S. Civil Service Commission
1900 E Street NW
Washington, DC 20415

1 C.S. WINIEWICZ
U.S. CIVIL SERVICE COMMISSION
REGIONAL PSYCHOLOGIST
230 S. DEARBORN STREET
CHICAGO, IL 60604

1 Dr. Joseph L. Young, Director
Memory & Cognitive Processes
National Science Foundation
Washington, DC 20550

Non Govt

1 PROF. EARL A. ALLUISI
DEPT. OF PSYCHOLOGY
CODE 287
OLD DOMINION UNIVERSITY
NORFOLK, VA 23508

1 1 psychological research unit
Dept. of Defense (Army Office)
Campbell Park Offices
Canberra ACT 2600, Australia

1 MR. SAMUEL BALL
EDUCATIONAL TESTING SERVICE
PRINCETON, NJ 08540

1 Dr. Nicholas A. Bond
Dept. of Psychology
Sacramento State College
600 Jay Street
Sacramento, CA 95819

1 Dr. Lyle Bourne
Department of Psychology
University of Colorado
Boulder, CO 80302

1 Dr. John Seeley Brown
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

1 Dr. John B. Carroll
Psychometric Lab
Univ. of No. Carolina
Davie Hall 013A
Chapel Hill, NC 27514

1 Dr. Kenneth E. Clark
College of Arts & Sciences
University of Rochester
River Campus Station
Rochester, NY 14627

1 Dr. Norman Cliff
Dept. of Psychology
Univ. of So. California
University Park
Los Angeles, CA 90007

1 Dr. Meredith Crawford
Department of Engineering Administrat
George Washington University
Suite 805
2101 L Street N. W.
Washington, DC 20037

1 DR. RENE V. DAWIS
DEPT. OF PSYCHOLOGY
UNIV. OF MINNESOTA
75 E. RIVER RD.
MINNEAPOLIS, MN 55455

1 Dr. Ruth Day
Center for Advanced Study
in Behavioral Sciences
202 Junipero Serra Blvd.
Stanford, CA 94305

1 Dr. Marvin D. Dunnette
N492 Elliott Hall
Dept. of Psychology
Univ. of Minnesota
Minneapolis, MN 55455

1 MAJOR [illegible] EVONIC
CANA[illegible]ORCES PERS. APPLIED RESEARCH
1107 [illegible] ROAD
TORONTO, ONTARIO, CANADA

1 Dr. Richard L. Ferguson
The American College Testing Program
P.O. Box 168
Iowa City, IA 52240

1 Dr. Victor Fields
Dept. of Psychology
Montgomery College
Rockville, MD 20850

1 Dr. Edwin A. Fleishman
Advanced Research Resources Organ.
8555 Sixteenth Street
Silver Spring, MD 20910

1 Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

1 DR. ROBERT GLASER
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 DR. JAMES G. GREENO
LRDC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Dr. Ron Hambleton
School of Education
University of Massechusetts
Amherst, MA 01002

1 Dr. Richard S. Hatch
Decision Systems Assoc., Inc.
350 Fortune Terrace
Rockville, MD 20854

2 Dr. Barbara Hayes-Roth
The Rand Corporation
1700 Main Street
Santa Monica, CA 90406

1 Dr. James R. Hoffman
Department of Psychology
University of Delaware
Newark, DE 19711

1 Library
HumRRO/Western Division
27857 Berwick Drive
Carmel, CA 93921

1 Dr. Earl Hunt
Dept. of Psychology
University of Washington
Seattle, WA 98105

1 Dr. Steven W. Keele
Dept. of Psychology
University of Oregon
Eugene, OR 97403

1 Dr. Walter Kintsch
Department of Psychology
University of Colorado
Boulder, CO 80302

1 Mr. Marlin Kroger
1117 Via Goleta
Palos Verdes Estates, CA 90274

1 LCOL. C.R.J. LAFLEUR
PERSONNEL APPLIED RESEARCH
NATIONAL DEFENSE HQS
101 COLONEL BY DRIVE
OTTAWA, CANADA K1A 0K2

1 Dr. Robert A. Levit
Manager, Behavioral Sciences
The BDM Corporation
7915 Jones Branch Drive
McClean, VA 22101

1 Dr. Frederick M. Lord
Educational Testing Service
Princeton, NJ 08540

1 Dr. Robert R. Mackie
Human Factors Research, Inc.
6780 Cortona Drive
Santa Barbara Research Pk.
Goleta, CA 93017

1 Dr. Richard B. Millward
Dept. of Psychology
Hunter Lab.
Brown University
Providence, RI 82912

1 Dr. Donald A Norman
Dept. of Psychology C-009
Univ. of California, San Diego
La Jolla, CA 92093

1 Dr. Melvin R. Novick
Iowa Testing Programs
University of Iowa
Iowa City, IA 52242

1 Dr. Jesse Orlansky
Institute for Defense Analysis
400 Army Navy Drive
Arlington, VA 22202

1 MR. LUIGI PETRULLO
2431 N. EDGEWOOD STREET
ARLINGTON, VA 22207

1 DR. STEVEN M. PINE
4950 Douglas Avenue
Golden Valley, MN 55416

1 DR. PETER POLSON
DEPT. OF PSYCHOLOGY
UNIVERSITY OF COLORADO
BOULDER, CO 80302

1 DR. DIANE M. RAMSEY-KLEE
R-K RESEARCH & SYSTEM DESIGN
3947 RIDGEMONT DRIVE
MALIBU, CA 90265

1 MIN. RET. M. RAUCH
P II 4
BUNDESMINISTERIUM DER VERTEIDIGUNG
POSTFACH 161
53 BONN 1, GERMANY

1 Dr. Mark D. Reckase
Educational Psychology Dept.
University of Missouri-Columbia
12 Hill Hall
Columbia, MO 65201

1 Dr. Joseph W. Rigney
Univ. of So. California
Behavioral Technology Labs
3717 South Hope Street
Los Angeles, CA 90007

1 Dr. Andrew M. Rose
American Institutes for Research
1055 Thomas Jefferson St. NW
Washington, DC 20007

1 Dr. Leonard L. Rosenbaum, Chairman
Department of Psychology
Montgomery College
Rockville, MD 20850

1 Dr. Ernst Z. Rothkopf
Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

1 PROF. FUMIKO SAMEJIMA
DEPT. OF PSYCHOLOGY
UNIVERSITY OF TENNESSEE
KNOXVILLE, TN 37916

1 DR. WALTER SCHNEIDER
DEPT. OF PSYCHOLOGY
UNIVERSITY OF ILLINOIS
CHAMPAIGN, IL 61820

1 DR. ROBERT J. SEIDEL
INSTRUCTIONAL TECHNOLOGY GROUP
HUMRRO
300 N. WASHINGTON ST.
ALEXANDRIA, VA 22314

1 Dr. Richard Snow
School of Education
Stanford University
Stanford, CA 94305

1 Dr. Robert Sternberg
Dept. of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

1 DR. ALBERT STEVENS
BOLT BERANEK & NEWMAN, INC.
50 MOULTON STREET
CAMBRIDGE, MA 02138

1 DR. PATRICK SUPPES
INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
STANFORD UNIVERSITY
STANFORD, CA 94305

1 Dr. Kikumi Tatsuoka
Computer Based Education Research Laboratory
252 Engineering Research Laboratory
University of Illinois
Urbana, IL 61801

1 DR. PERRY THORNDYKE
THE RAND CORPORATION
1700 MAIN STREET
SANTA MONICA, CA 90406

1 Dr. Benton J. Underwood
Dept. of Psychology
Northwestern University
Evanston, IL 60201

1 Dr. Claire E. Weinstein
Educational Psychology Dept.
Univ. of Texas at Austin
Austin, TX 78712

1 DR. SUSAN E. WHITELY
PSYCHOLOGY DEPARTMENT
UNIVERSITY OF KANSAS
LAWRENCE, KANSAS 66044

PREVIOUS PUBLICATIONS

Proceedings of the 1977 Computerized Adaptive Testing Conference. July 1978.

Research Reports

78-2. The Effects of Knowledge of Results and Test Difficulty on Ability Test Performance and Psychological Reactions to Testing. September 1978.

78-1. A Comparison of the Fairness of Adaptive and Conventional Testing Strategies. August 1978. (NTIS No. AD A059436)

77-7. An Information Comparison of Conventional and Adaptive Tests in the Measurement of Classroom Achievement. October 1977. (NTIS No. AD A047495)

77-6. An Adaptive Testing Strategy for Achievement Test Batteries. October 1977. (NTIS No. AD A046062)

77-5. Calibration of an Item Pool for the Adaptive Measurement of Achievement. September 1977. (NTIS No. AD A044828)

77-4. A Rapid Item-Search Procedure for Bayesian Adaptive Testing. May 1977 (NTIS No. AD A041090)

77-3. Accuracy of Perceived Test-Item Difficulties. May 1977. (NTIS No. AD A041084)

77-2. A Comparison of Information Functions of Multiple-Choice and Free-Response Vocabulary Items. April 1977.

77-1. Applications of Computerized Adaptive Testing. March 1977. (NTIS No. AD A038114)

Final Report: Computerized Ability Testing, 1972-1975. April 1976. (NTIS No. AD A024516)

76-5. Effects of Item Characteristics on Test Fairness. December 1976. (NTIS No. AD A035393)

76-4. Psychological Effects of Immediate Knowledge of Results and Adaptive Ability Testing. June 1976. (NTIS No. AD A027170)

76-3. Effects of Immediate Knowledge of Results and Adaptive Testing on Ability Test Performance. June 1976. (NTIS No. AD A027147)

76-2. Effects of Time Limits on Test-Taking Behavior. April 1976. (NTIS No. AD A024422)

76-1. Some Properties of a Bayesian Adaptive Ability Testing Strategy. March 1976. (NTIS No. AD A022964)

75-6. A Simulation Study of Stradaptive Ability Testing. December 1975. (NTIS No. AD A020961)

75-5. Computerized Adaptive Trait Measurement: Problems and Prospects. November 1975. (NTIS No. AD A018675)

75-4. A Study of Computer-Administered Stradaptive Ability Testing. October 1975. (NTIS No. AD A018758)

75-3. Empirical and Simulation Studies of Flexilevel Ability Testing. July 1975. (NTIS No. AD A013185)

75-2. TETREST: A FORTRAN IV Program for Calculating Tetrachoric Correlations. March 1975. (NTIS No. AD A007572)

75-1. An Empirical Comparison of Two-Stage and Pyramidal Adaptive Ability Testing. February 1975. (NTIS No. AD A006733)

74-5. Strategies of Adaptive Ability Measurement. December 1974. (NTIS No. AD A004270)

74-4. Simulation Studies of Two-Stage Ability Testing. October 1974. (NTIS No. AD A001230)

74-3. An Empirical Investigation of Computer-Administered Pyramidal Ability Testing. July 1974. (NTIS No. AD 783553)

74-2. A Word Knowledge Item Pool for Adaptive Ability Measurement. June 1974. (NTIS No. AD 781894)

74-1. A Computer Software System for Adaptive Ability Measurement. January 1974. (NTIS No. AD 773961)

73-4. An Empirical Study of Computer-Administered Two-Stage Ability Testing. October 1973. (NTIS No. AD 768993)

73-3. The Stratified Adaptive Computerized Ability Test. September 1973. (NTIS No. AD 768376)

73-2. Comparison of Four Empirical Item Scoring Procedures. August 1973.

73-1. Ability Measurement: Conventional or Adaptive? February 1973. (NTIS No. AD 757788).

AD Numbers are those assigned by the Defense Documentation Center, for retrieval through the National Technical Information Service.

Copies of these reports are available, while supplies last, from:

Psychometric Methods Program
Department of Psychology
N660 Elliott Hall
University of Minnesota
75 East River Road
Minneapolis, Minnesota 55455